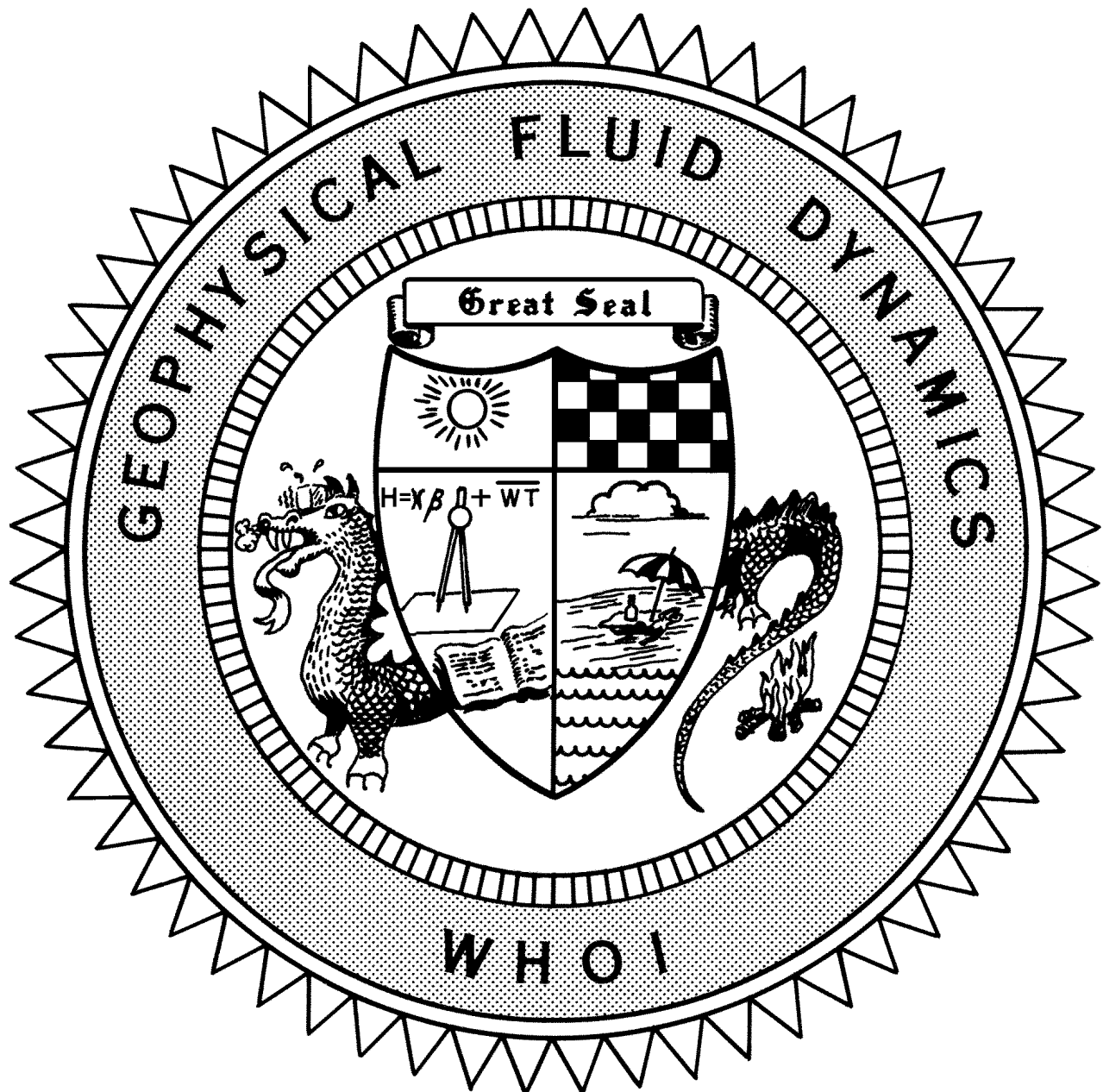


***2008 Program of Study:
Perspectives and Challenges in GFD***



**Course Lectures
Fellows Project Reports**

WHOI-2009-02

**2008 Program of Studies:
Perspectives and Challenges in GFD**

by

Neil J. Balmforth and George Veronis, Co-Directors

Stephen W. Childress, Charles Doering, Kerry Emanuel, Stephan Fauve, Raffaele Ferrari,
Christopher J.R. Garrett, Herbert E. Huppert, John Marshall, Raymond T. Pierrehumbert
and Timour Radko, Co-Principal Lecturers

March 2009

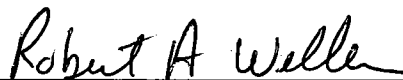
Technical Report

Funding was provided by the National Science Foundation under grant OCE-0325296 and
by the Office of Naval Research, Processes and Prediction Division, Physical
Oceanography Program under grant N00014-07-10776.

Reproduction in whole or in part is permitted for any purpose of the United
States Government. This report should be cited as Woods Hole Oceanog. Inst.
Tech. Rept., WHOI-2009-02.

Approved for public release; distribution unlimited.

Approved for Distribution:



Robert A. Weller, Chair

Department of Physical Oceanography

Preface

N. J. Balmforth and G. Veronis

22nd October 2008

The GFD Program in 2008 ran with a special theme, in view of the 50th anniversary of the beginning of the program. The first two weeks of Principal Lectures dealt with "Perspectives and Challenges" of the subject, and we commemorated by having ten different lecturers with a diversity of viewpoints and opinions. Though the lectures thereby took less of a pedagogical flavour, each afternoon we saw the fellows grill the lecturers in informal afternoon tutorials.

Also by way of celebration, the Oceanographic Institution threw a 50th anniversary party in honour of the Program and the "Founding Fathers." This was held on one of the lawns in front of Walsh Cottage on Friday, June 27th. We saw a variety of familiar faces, some from the very first years of the Program, others from more recent summers. George Veronis and Jack Whitehead read short pieces about the past fifty years, including the unveiling of the "2008 AGU Award for Excellence in Geophysical Education" awarded to the Program by the American Geophysical Union. The party was skillfully organized by Janet Fields and uniformly enjoyed by all.

In 2008 the Sears Public Lecture was delivered by Andy Ingersoll (CalTech) and entitled "Weather and Climates of Other Worlds: Lessons for Earth." Andy, a participant of the program for many years, described the insights that could be drawn regarding our own climate from observing those of the other planets in the solar system, observations made particularly clear and compelling by recent space missions. Over a hundred people gathered at Redfield for the lecture, many of whom paused afterwards to enjoy conversation and refreshments at the reception that followed.

Other novelties at the 2008 program included Bill Dewar's delivery of a seminar at Walsh in full Scottish regalia, kilt and all. Last, and of course most importantly, we had an industrious group of fellows who rose to the challenge of the Program and contributed tremendously to the anniversary. As always, the WHOI Academic Programs Office and the Physical Oceanography Department provided the administrative backbone to the summer, and Walsh Cottage was perfectly unchanged in its intimacy and rustic atmosphere. Jeanne Fleming, Penny Foster and Janet Fields contributed importantly to the smooth running of the program.

Contents

PREFACE	i
CONTENTS	ii
PARTICIPANTS	iv
GFD PHOTOGRAPHS	vii
LECTURE SCHEDULE	ix
PRINCIPAL LECTURES	1
Steve Childress: <i>Dynamo Theory and GFD</i>	2
Charlie Doering: <i>Convection, Stability and Turbulence</i>	19
Kerry Emanuel: <i>Waves and Vortices Driven by Interfacial Fluxes</i>	38
Stephan Fauve: <i>Turbulent Dynamos: experiments, nonlinear saturation of the magnetic field and field reversals</i>	54
Raffaele Ferrari: <i>Energetics of a Turbulent Ocean</i>	70
Chris Garrett: <i>Ocean Mixing, Internal Tides and Tidal Power</i>	82
Herbert Huppert: <i>Gravity Currents and Solidification</i>	96
John Marshall: <i>GFD Experiments in Climate</i>	117
Ray Pierrehumbert: <i>Atmospheric Escape</i>	133
Timour Radko: <i>Double-Diffusive Convection</i>	143
FELLOWS REPORTS	162
Sylvain Barbot: <i>Models of volcanic tremor and singing icebergs</i>	163
Chris Cawthorn: <i>First Contact in a Viscous Fluid</i>	188
Christophe Gissinger: <i>Energy and dissipation in MHD systems</i>	215
Céline Guervilly: <i>MHD Toy Model of the Solar Radiative Zone</i>	245

George Hagstrom: <i>Bounds on Surface Stress Driven Flows</i>	256
Ian Hewitt: <i>Continual Skipping - Getting a Kick From Water</i>	274
Malte Jansen: <i>On the Organization of Geostrophic Circulations Over Large Scale Topography by Eddy-Diffusion of PV</i>	303
Amrita Shravat: <i>Turbulent Mixing in Shear Driven Stratified Fluid</i>	318
Toby Wood: <i>Crumpling of a Thin Ice Sheet Due to Incident Flow</i>	336
Yutian Wu: <i>Equilibria of Diffusive Moist Static Energy Balance Model . .</i>	348

2008 GFD Fellows, Staff and Visitors

The Fellows

Sylvain D. Barbot	University of California at San Diego
Christopher J. Cawthorn	University of Cambridge
Christophe J. P. Gissinger	Ecole Normale Supérieure
Céline Guervilly	Université de Grenoble
George I. Hagstrom	University of Texas at Austin
Ian J. Hewitt	University of Oxford
Malte F. Jansen	Massachusetts Institute of Technology
Amrita Shrivastava	Oxford University
Toby S. Wood	University of Cambridge
Yutian Wu	Columbia University

Staff and Visitors

James Anderson	Stevens Institute of Technology
Sheekela Baker-Yeboah	Massachusetts Institute of Technology
Neil J. Balmforth	University of British Columbia
Andrew Belmonte	Pennsylvania State University
Onno Bokhove	University of Twente
Michael Brenner	Harvard University
Gregory Buck	Saint Anselm College
Oliver Buhler	New York University
John Bush	Massachusetts Institute of Technology
Jeff Carpenter	University of British Columbia
Colm-cille P. Caufield	University of Cambridge
Claudia Cenedese	Woods Hole Oceanographic Institution
Paola Cessi	University of California, San Diego
Eric P. Chassignet	Florida State University
Stephen W. Childress	New York University
Greg Chini	University of New Hampshire
Aline Cotel	University of Michigan
Predrag Cvitanovic	Georgia Institute of Technology
Stuart Dalziel	University of Cambridge
William K. Dewar	Florida State University
Charles Doering	University of Michigan
Kerry Emanuel	Massachusetts Institute of Technology
Stephan Fauve	Ecole Normale Supérieure
Alexy V. Federov	Yale University
Raffaele Ferrari	Massachusetts Institute of Technology
Glenn Flierl	Massachusetts Institute of Technology

Basile Gallet	Ecole Normale Supérieure
Pascale Garaud	University of California, Santa Cruz
Christopher J. R. Garrett	University of Victoria
John Gibson	Georgia Institute of Technology
Gary Glatzmaier	University of California, Santa Cruz
Jerry Gollub	Haverford College
Jeroen Hazewinkel	University of Cambridge
Karl R. Helfrich	Woods Hole Oceanographic Institution
Rainer Hollerbach	University of Leeds
Louis N. Howard	Massachusetts Institute of Technology
Herbert E. Huppert	University of Cambridge
Andrew P. Ingersoll	California Institute of Technology
Edward Johnson	University College
Bror Jonsson	Boston University
Joseph B. Keller	Stanford University
Patrice Klein	IFREMER
Arshad Kudroli	Clark University
Joseph LaCasce	The Norwegian Meteorological Institute
Frederic B. Laliberte	New York University
Norman R. Lebovitz	University of Chicago
Liming Li	Cornell University
Zhi Lin	University of Michigan
Jennifer Mackinnon	University of California, San Diego
Amala Mahadevan	Boston University
Lakshminarayanan Mahadevan	Harvard University
Filippo Maimone	Aeronautica Militare
Willem V.R. Malkus	Massachusetts Institute of Technology
John Marshall	Massachusetts Institute of Technology
James McElwaine	University of Cambridge
Philip J. Morrison	University of Texas at Austin
Walter Munk	University of California, San Diego
Johan Nilsson	Stockholm University
Francesco Paparella	University of Lecce
Thomas Peacock	Massachusetts Institute of Technology
Joseph Pedlosky	Woods Hole Oceanographic Institution
W. Richard Peltier	University of Toronto
Raymond T. Pierrehumbert	University of Chicago
Lawrence J. Pratt	Woods Hole Oceanographic Institution
Antonello Provenzale	Istituto di Scienze Dell'Atmosfera
Timour Radko	Naval Postgraduate School
Alan W. Rempel	University of Oregon
Alison C. Rust	University of Bristol

Roger M. Samelson	Oregon State University
Tiffany A. Shaw	University of Toronto
Vitalii Sheremet	University of Rhode Island
Alexander V. Soloviev	Nova Southeastern University
Edward A. Spiegel	Columbia University
Ravi Srinivasan	Brown University
Bruce R. Sutherland	University of Alberta
Lynne Talley	University of California, San Diego
Jean-Luc Thiffeault	University of Wisconsin
Mary-Louise E. Timmermanns	Woods Hole Oceanographic Institution
Adrienne Traxler	University of California, Santa Cruz
Eli Tziperman	Harvard University
Geoffrey K. Vallis	Princeton University
George Veronis	Yale University
John A. Whitehead	Woods Hole Oceanographic Institution
Andrew Woods	University of Cambridge
Carl Wunsch	Massachusetts Institute of Technology
William R. Young	University of California, San Diego
Jun Zhang	New York University
Xiaoqian Zhang	Texas A&M University



Top row, from left to right: Penny Foster, Xiaoqian Zhang; Colm-cille Caulfield; Michael Brenner; John Taylor; Ted Johnson; Filippo Maimone; (unknown); Adrienne Traxler; Jennifer MacKinnon; Jeroen Hazewinkel; John Marshall; Ray Pierrehumbert; Raffaele Ferrari; Bill Dewar; Lou Howard

Middle row, from left to right: Ed Spiegel (standing); Charles Doering; Jean-Luc Thiffeault; L. Mahadevan; Herbert Huppert; Joe Keller; George Veronis; Larry Pratt; Eric Chassignet; Steve Childress; Pascale Garaud; Lucia Bunge; Johan Nilsson; Neil Balmforth; Norman Lebovitz (standing); Timour Radko (standing); Amala Mahadevan (standing)

Bottom row, from left to right: Ian Hewitt, Christophe Gissinger, Toby Wood, Malte Jansen, Sylvain Barbot, George Hagstrom, Christopher Cawthorn, Celine Guervilly, Amrita Shravat, Yutian Wu

GFD 2008



Raymond T. Pierrehumbert



Charles Doering



Christopher J.R. Garrett



Timour Radko



Raffaele Ferrari



John Marshall



Kerry Emanuel



Stephen W. Childress



Stephan Fauve



Herbert E. Huppert

GFD 2008 Seminar Schedule

Week 1:

Monday, June 16	Steve Childress	Dynamo theory and GFD
Tuesday, June 17	Stefan Fauve	Turbulent dynamos
Wednesday, June 18	Charles Doering	Convection, stability and turbulence
Thursday, June 19	Christopher Garrett	Ocean mixing and tidal power
Friday, June 20	Kerry Emanuel	Waves and vortices driven by interfacial fluxes

Week 2:

Monday, June 23	Raffaele Ferrari	The oceanic energy cycle
Tuesday, June 24	Raymond Pierrehumbert	Atmospheric escape
Wednesday, June 25	Herbert Huppert	Geological fluid mechanics
Thursday, June 26	John Marshall	GFD experiments in climate
Friday, June 27	Timour Radko	Double-diffusive convection

Week 3:

Monday, June 30	Jean-Luc Thiffeault	Mixing hits a wall
	Colm-cille Caulfield	Feeling the pinch: Time-dependent plume dynamics
Tuesday, July 1	Antonello Provenzale	Rain, droughts & veggies: A simple model of soil-vegetation atmosphere dynamics
Wednesday, July 2	Eric Chassignet	Ocean modeling in quasi-Lagrangian vertical coordinates
	Liming Li	Waves and jets on the four giant planets
Thursday, July 3	Jim McElwaine	Washboard roads

Week 4:

Monday, July 7	William Dewar	Inviscid dissipation of balanced flow
	Alexander Soloviev	Hydrodynamics and remote sensing of far wakes of ships
Tuesday, July 8	Dick Peltier	PV staircases and the dynamics of Jupiter's atmosphere
Wednesday, July 9	Larry Pratt	Circulation and exchange in marginal seas
Thursday, July 10	L. Mahadevan	Soft lubrication and adhesion
	Amala Mahadevan	GFD and Biology
Friday, July 11	Ted Johnson	Vortex patches near boundaries

Week 5:

Monday, July 14	Claudia Cenedese	A new entrainment parameterization for mixing in overflows
Tuesday, July 15	Rainer Hollerbach	Instabilities of Stewartson layers and Taylor columns
Wednesday, July 16	Jeff Carpenter	Holmboe's instability
Thursday, July 17	Phil Morrison	Gyroviscosity and magnetofluid models
	Francesco Paparella	Granular gases, inelastic collapse and bouncing balls
Friday, July 18	Michael Brenner	Splashing and splitting

Week 6:

Monday, July 21	Andrew Belmonte	Gel fingers
Tuesday, July 22	Andrew Woods	Fluid dynamics of carbon sequestration
Wednesday, July 23	Stuart Dalziel	Maximal mixing efficiency
Thursday, July 24	Tiffany Shaw	Wave-activity conservation laws and their application to subgrid-scale parameterization in climate models

Week 7:

Monday, July 28	George Veronis	Effect of double diffusion on the dam break experiment
Tuesday, July 29	Alan Rempel	Seepage flows and glacier sliding
	Andrew Fowler	Mantle convection
Wednesday, July 30	Alexey Federov	Energy transfer from the winds to the thermocline on ENSO timescales
	John Bush	The fluid trampoline: Droplets bouncing on a soap film
Thursday, July 31	Jerry Gollub	Tracking topological features to characterize chaotic flow
Friday, August 1	Joe Keller	Some problems in solid and fluid mechanics

Week 8:

Monday, August 4	Pedlosky Symposium	
Tuesday, August 5	Bruce Sutherland	The evolution of finite-amplitude internal gravity waves
	Predrag Cvitanovic	Geometry of boundary shear turbulence: A stroll through 61,506 Dimensions
Wednesday, August 6	Aline Cotel	New results on turbulent entrainment in stratified flows
	Oliver Buhler	Propagation and saturation of internal tides
	Andrew Ingersoll	Weather and climates of other worlds: Lessons for Earth
Thursday, August 7	Gary Glatzmaier	Effects of rotation and density stratification on the interior dynamics of giant planets
	Onno Bokhove	Magma-rock interactions in gelatin-water laboratory experiments
Friday, August 8	Walter Munk	An embarrassing sea truth; steepness, spread and skewness of ocean waves
	Carl Wunsch	Ocean variability, memory and trends

Week 10:

Tuesday, August 19	Chris Cawthorn	Close encounters of the viscous kind: first contact in a viscous fluid
	Sylvain Barbot	Models of volcanic tremors & singing icebergs
	Céline Guervilly	Toy model of the solar radiative zone
Wednesday, August 20	George Hagstrom	Bounds for shear-stress driven flows
	Malte Jansen	On the interaction of eddies with large-scale topography
	Christophe Gissinger	Energy and dissipation in MHD systems
	Amrita Shravat	Experiment on mixing induced by a horizontal disc
Thursday, August 21	Yutian Wu	Equilibria of diffusive moist static energy balance models
	Toby Wood	(Huge) problems with surface tension
	Ian Hewitt	Continual skipping on water

PRINCIPAL LECTURES

Dynamo theory and GFD

Steve Childress

16 June 2008

1 Origins of the dynamo theory

Dynamo theory studies a conducting fluid moving in a magnetic field; the motion of the body through the field acts to generate new magnetic field, and the system is called a dynamo if the magnetic field so produced is self-sustaining.

1.1 Early ideas point the way

In the distant past, there was the idea that the earth was a permanent magnet. In the 1830s Gauss analyzed the structure of the Earth's magnetic field using potential theory, decomposing the field into harmonics. The strength of the dominant field was later found to change with time.

In 1919 Sir Joseph Larmor drew on the induction of currents in a moving conductor, to suggest that sunspots are maintained by magnetic dynamo action. P. M. Blackett proposed that magnetic fields should be produced by the rotation of fluid bodies.

The 'current consensus' is that the Earth's magnetic field is the result of a regenerating dynamo action in the fluid core. The mechanism of generation of the field is closely linked dynamically with the rotation of the Earth. Similar ideas are believed to apply to the solar magnetic field, to other planetary fields, and perhaps to the magnetic field permeating the cosmos.

1.2 Properties of the Earth and its Magnetic Field

The magnetic field observed at the Earth's surface changes polarity irregularly. The non-dipole components of the surface field also vary with time over many time scales greater than decades, and have a persistent drift to the west. The fluid core of the Earth is a spherical annulus, bounded by the solid inner core and the mantle. It is believed that the motion of the inner and outer core are sufficient to drive the geodynamo.

2 The homogeneous kinematic dynamo

The pre-Maxwell equations for a homogeneous moving conductor are

$$\begin{aligned}
\nabla \times \mathbf{B} &= \mu \mathbf{J} \\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\
\mathbf{J} &= \sigma (\mathbf{E} + \mathbf{u} \times \mathbf{B}) \\
\nabla \cdot \mathbf{B} &= 0 \\
\nabla \cdot \mathbf{E} &= q/\epsilon
\end{aligned} \tag{1}$$

Combined, these equations simplify to

$$\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{u} \times \mathbf{B}) - \eta \nabla^2 \mathbf{B} = 0 \tag{2}$$

where $\eta = (\mu\sigma)^{-1}$ is the magnetic diffusivity. This equation is called the magnetic induction equation.

2.1 Kinematic Dynamo Model of the Earth's Core

Neglecting the inner core, the conducting fluid is contained in a sphere of radius $r = r_c$. The exterior is regarded as free space. The equations are to be solved with a prescribed divergence-free (the core fluid is assumed incompressible) velocity field which is independent of time. On $r = r_c$, the magnetic field is continuous owing to the absence of magnetic monopoles and a concentrated surface current layer. The magnetic field on $r > r_c$ also matches with an external vacuum magnetic field

$$\mathbf{B}_e = \nabla \phi_e. \tag{3}$$

The tangential component of

$$\mathbf{E} = \eta \nabla \times \mathbf{B} - \mathbf{u} \times \mathbf{B} \tag{4}$$

is also continuous at $r = r_c$. We take the external field to decay like a potential dipole at $r = \infty$, i.e., $\mathbf{B} \approx O(r^{-3})$.

We pass to a dimensionless form using a characteristic length scale L , velocity scale U and time scale L/U . Using the vector identity

$$\nabla \times (\mathbf{u} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \mathbf{u} + \mathbf{u} \nabla \cdot \mathbf{B} - \mathbf{u} \cdot \nabla \mathbf{B} - \mathbf{B} \nabla \cdot \mathbf{u}, \tag{5}$$

and since the velocity and magnetic fields are divergence free ($\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{B} = 0$), the induction equation can be written in the form

$$\frac{D\mathbf{B}}{Dt} - \frac{1}{R} \nabla^2 \mathbf{B} = \mathbf{B} \cdot \nabla \mathbf{u} \tag{6}$$

in terms of the material derivative

$$\frac{D\mathbf{B}}{Dt} = \frac{\partial \mathbf{B}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{B}. \tag{7}$$

Here R is the *magnetic Reynolds number*, $R = UL/\eta = UL\sigma\mu$. For $R \ll 1$, the magnetic field diffuses easily through the conductor. For $R \gg 1$, the magnetic field is ‘frozen into the moving conductor’. In this case, distortion and stretching of the field lines are caused by the term $\mathbf{B} \cdot \nabla \mathbf{u}$. Notice that eq. (6) corresponds to the vorticity equation upon substitution of $\nabla \times \mathbf{u}$ in place of \mathbf{B} .

2.2 Kinematic Dynamo as an Eigenvalue Problem

For a given velocity field \mathbf{u} , eq. (6), is linear for the magnetic field \mathbf{B} , and we may separate variables

$$\begin{aligned}\mathbf{B} &= e^{\lambda t} \mathbf{b}(\mathbf{x}), & r < 1 \\ \mathbf{B} &= e^{\lambda t} \nabla \phi, & r > 1\end{aligned}\tag{8}$$

Upon change of variable (8), the induction equation becomes

$$\mathcal{L}\mathbf{b} = \frac{1}{R} \nabla^2 \mathbf{b} - \mathbf{u} \cdot \nabla \mathbf{b} + \mathbf{b} \cdot \nabla \mathbf{u} = \lambda \mathbf{b}\tag{9}$$

Here, length dimension L is $L = r_c$. The values of λ allowing acceptable solutions \mathbf{b} are the eigenvalues, and depend on R . We say $\mathbf{u}(\mathbf{x})$ is a (steady) kinematic dynamo if for some $R > 0$ there exists an eigenvalue $\lambda(R)$ such that the real part of λ is greater than zero. Elsasser, Bullard, and others developed a theory of the spherical dynamo in this setting.

Note that although the kinematic equation is linear in \mathbf{B} , the dynamo problem has a nonlinear character, since the ‘correct’ velocity field \mathbf{u} is not known *a priori*. Mathematically, we are trying to find the right function $\mathbf{u}(\mathbf{x})$ determining \mathcal{L} , so that the dynamo property is realized. More generally, we might consider a time periodic \mathbf{x} and an analogous Floquet problem.

2.3 Expansion in eigenfunctions

Consider the space of complex-valued divergence-free vector fields $\mathbf{b}(\mathbf{x})$ in $r < 1$, matching continuously with a potential field in $r > 1$. The eigenfunctions \mathbf{b}_n satisfy

$$\begin{aligned}\mathcal{L}\mathbf{b}_n &= \lambda_n \mathbf{b}_n, & r < 1 \\ \mathbf{b}_n &= \nabla \phi_n, & r > 1\end{aligned}\tag{10}$$

and are continuous on $r = 1$. The eigenvalues are known to be discrete and countable, and the eigenfunctions are complete in the above space. If the eigenfunctions were known to be mutually orthogonal, in the inner product

$$(\mathbf{b}_m, \mathbf{b}_n) = \int_{\mathcal{R}^3} \mathbf{b}_m^* \mathbf{b}_n dV\tag{11}$$

then, we would have

$$\int_{\mathcal{R}^3} |\mathbf{B}|^2 = \sum_{n=1}^{\infty} e^{2\Re(\lambda_n t)} \|\mathbf{b}_n\|^2\tag{12}$$

Thus, if $\Re(\lambda_n) \neq 0$ for all n the magnetic energy would be bounded by its initial value. This is the case for a *normal* operator on a Hilbert space, one which commutes with its adjoint.

However, the operator \mathcal{L} is not normal so the energy is not a simple sum of decaying non-negative terms. In fact, substantial transient growth of field energy is possible even if \mathbf{u} is not a dynamo. This complicates numerical proofs of the dynamo property. The dynamo problem is one of the best examples of non-normality, and the effects become especially pronounced in the limit of large magnetic Reynolds number R . In the case of the Earth, R is of the order 10^3 .

2.4 Free-decay modes for a rigid spherical core

It can be shown that in \mathcal{R}^3 , the following decomposition of a divergence-free vector field \mathbf{A} is possible

$$\mathbf{A} = \nabla \times (T \mathbf{r}) + \nabla \times \nabla \times (P \mathbf{r}) \quad (13)$$

where $T = T(\mathbf{r})$ and $P = P(\mathbf{r})$ are scalar functions called the toroidal and poloidal components of \mathbf{A} , respectively (figure 1). Consider this decomposition for $\mathbf{u} = 0$, so that we have

$$\frac{1}{R} \nabla^2 \mathbf{b}_n = \lambda_n \mathbf{b}_n \quad (14)$$

The eigenfunctions \mathbf{b}_n are decomposed as follows:

$$\mathbf{b}_n = \nabla \times (T_n \mathbf{r}) + \nabla \times \nabla \times (P_n \mathbf{r}) \quad (15)$$

In the case of an axis-symmetric problem, one can express the eigenfunctions in terms of their components in spherical coordinates as follows:

$$\mathbf{b}_n = \frac{1}{r} L_2 P \mathbf{e}_r + \frac{1}{r} \frac{\partial}{\partial \theta} \frac{\partial}{\partial r} r P \mathbf{e}_\theta + \frac{\partial}{\partial \theta} T \mathbf{e}_\phi \quad (16)$$

where L_2 is the surface Laplacian. If \mathbf{b}_n is expanded in spherical harmonics

$$\mathbf{b}_n = \sum_{l=1}^{\infty} \sum_{k=-l}^l b_l^m f(r) Y_l^m(\theta) e^{im\phi} \quad (17)$$

the surface operator has the property

$$L_2 \mathbf{b}_n = -l(l+1) \mathbf{b}_n \quad (18)$$

The exterior of the dynamo is current free, and with the pre-Maxwell equation $\nabla \times \mathbf{B} = \mu \mathbf{J}$ one obtains $\nabla \times \mathbf{B} = 0$. In terms of eigenfunctions, the curl of the magnetic field can be written

$$\begin{aligned} \nabla \times \mathbf{b}_n &= \nabla \times (-\nabla^2 (P_n \mathbf{r}) + \nabla \nabla \cdot (P_n \mathbf{r})) + \nabla \times \nabla \times (T_n \mathbf{r}) \\ &= \nabla \times (-\nabla^2 (P_n \mathbf{r})) + \nabla \times \nabla \times (T_n \mathbf{r}) \end{aligned} \quad (19)$$

where we have used the identity $\nabla \times \nabla \times \mathbf{A} = \nabla \nabla \cdot \mathbf{A} - \nabla^2 \mathbf{A}$. We reckon that the toroidal component of the magnetic field is the poloidal component of the curl of the field, and similarly $-\nabla^2 P_n$ becomes the toroidal component of the curl of the field:

$$\begin{aligned} (\nabla \times \mathbf{b}_n)_T &= -\nabla^2 P_n \\ (\nabla \times \mathbf{b}_n)_P &= T_n \end{aligned} \quad (20)$$

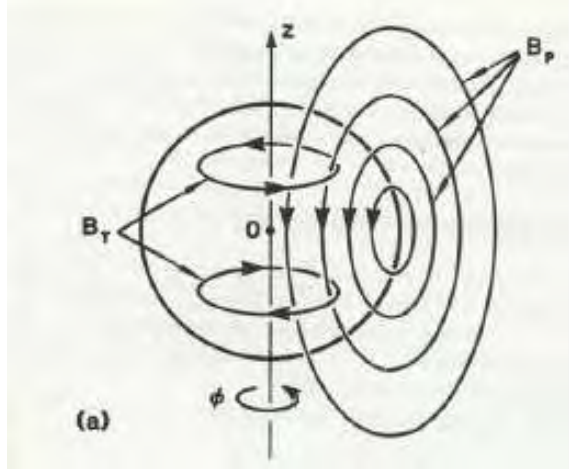


Figure 1: Toroidal and poloidal field lines for a sphere rotating about the z axis.

Using eq. (16) and considering the poloidal and toroidal components of the curl of the field, we obtain

$$\frac{1}{r}L_2T_n = -\frac{1}{r}l(l+1)T_n = 0 \quad (21)$$

so the toroidal field is zero at the exterior ($T_n = 0$, $r > 1$). Looking at the ϕ -component of the field rotational, we have

$$\frac{\partial}{\partial \theta}(\nabla^2 P_n) = 0 \quad (22)$$

Using the matching condition on the core boundary, one obtains $\nabla^2 P_n = 0$, $r > 1$.

It then follows that we may consider separately problems for T_n and P_n

$$\begin{aligned} \nabla^2 T_n &= \lambda_n T_n \\ \nabla^2 P_n &= \lambda_n P_n \\ T_n &= 0 \\ \nabla^2 P_n &= 0 \end{aligned} \quad \begin{aligned} r &< 1 \\ r &\geq 1 \end{aligned} \quad (23)$$

Here T_n , P_n , $\partial P_n / \partial r$ are continuous on $r = 1$, and $P_n \sim O(r^{-2})$ as $r \rightarrow \infty$.

The first few solutions are

$$T_1 = \frac{\sin \pi r}{r}, \quad \lambda_1 = -\pi^2, \quad (24)$$

where $r = |\mathbf{x}|$, (this mode is not believed to be relevant to the Earth's toroidal field except perhaps during reversals);

$$P_1 = \left(\frac{\pi \cos \pi r}{r} - \frac{\sin \pi r}{r^2} \right) \cos \theta, \quad r < 1, \quad P_1 = -\frac{\pi \cos \theta}{r^2}, \quad r > 1, \quad (25)$$

(this is the basic dipole component of the Earth's field with axis aligned with the rotation axis);

$$T_2 = \left(\frac{\mu \cos \mu r}{r} - \frac{\sin \mu r}{r^2} \right) \cos \theta, \quad \lambda_2 = -\mu^2 \approx -20.2, \quad (26)$$

where μ is the first zero of the spherical Bessel function j_1 , given by $\tan \mu = \mu$ (this is a plausible dominant toroidal component of the Earth's field).

2.5 The Omega Effect

An example of transient growth of magnetic energy in a non-dynamo is the so called *Omega Effect*, in which differential rotation causes deformation of poloidal field lines, which induce a toroidal field (figure 2). Let $\mathbf{u} = U(r)\mathbf{e}_\phi$, where $r = \sqrt{x^2 + y^2}$ in cylindrical polars, and represent the magnetic field, taken as symmetric with respect to the z -axis, in the alternative poloidal-toroidal decomposition

$$\mathbf{B} = B(r, z, t)\mathbf{e}_\phi + \nabla \times A(r, z, t)\mathbf{e}_\phi. \quad (27)$$

Now assume $A = rP$ determines the z -aligned dipole field and is independent of time. Taking the \mathbf{e}_ϕ components of the induction equation, B is then found to satisfy

$$\frac{\partial B}{\partial t} + r \frac{\partial A}{\partial z} \frac{\partial}{\partial r} \left(\frac{U(r)}{r} \right) - \frac{1}{R} \left(\nabla^2 - \frac{1}{r^2} \right) B = 0. \quad (28)$$

The second term here represents the omega effect - depending on $U(r)$ it can act as a source term for the toroidal component B , and shows how the poloidal field is coupled to the toroidal field and can cause the latter to grow. This however is not enough to create a dynamo, as we shall see below. Note a velocity $U(r)\mathbf{e}_\phi$ is easily realized in a rotating sphere of fluid as a *geostrophic flow*.

3 Establishing the possibility of a homogeneous dynamo

3.1 A negative result: Cowling's theorem

Cowling (1934) argued that no homogeneous dynamo can exist for axisymmetric \mathbf{u} and \mathbf{B} . This is a result of the failure of the poloidal field to be maintained against ohmic dissipation. Taking

$$\mathbf{u} = U(r, z, t)\mathbf{e}_\phi + \nabla \times \psi(r, z, t)\mathbf{e}_\phi, \quad \mathbf{B} = B(r, z, t)\mathbf{e}_\phi + \nabla \times A(r, z, t)\mathbf{e}_\phi, \quad (29)$$

we find

$$\frac{\partial A}{\partial t} - \frac{\partial \psi}{\partial z} \frac{1}{r} \frac{\partial}{\partial r} (rA) + \frac{\partial A}{\partial z} \frac{1}{r} \frac{\partial}{\partial r} (r\psi) - \frac{1}{R} \left(\nabla^2 - \frac{1}{r^2} \right) A = 0, \quad (30)$$

$$\frac{\partial B}{\partial t} + r \frac{\partial A}{\partial z} \frac{\partial}{\partial r} \left(\frac{U}{r} \right) - \frac{1}{r} \frac{\partial}{\partial r} (rA) \frac{\partial U}{\partial z} - r \frac{\partial \psi}{\partial z} \frac{\partial}{\partial r} \left(\frac{B}{r} \right) + \frac{1}{r} \frac{\partial}{\partial r} (r\psi) \frac{\partial B}{\partial z} - \frac{1}{R} \left(\nabla^2 - \frac{1}{r^2} \right) B = 0. \quad (31)$$

It is immediately clear from the equation for A that there is no coupling from the toroidal field to the poloidal field, and that the poloidal field will inevitably decay. Multiplying this equation for A by $r^2 A$ and integrating by parts, one finds

$$\frac{d}{dt} \int_{r < 1} \frac{1}{2} (rA)^2 dV = - \frac{1}{R} \int_{R^2} |\nabla(rA)|^2 dV \leq 0, \quad (32)$$

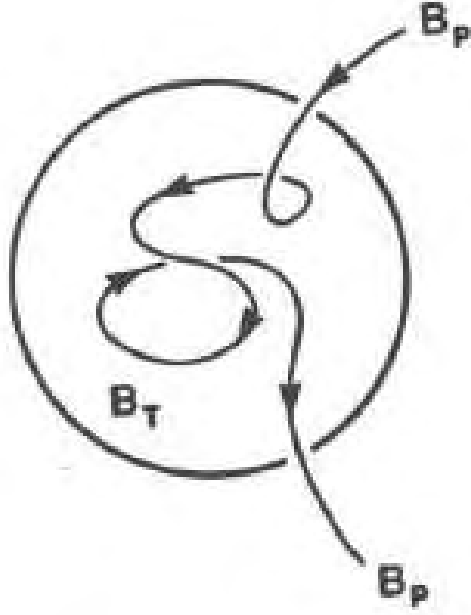


Figure 2: The Ω effect: differential rotation causes the twisting of poloidal field lines which generates toroidal field.

and this implies that the poloidal field must necessarily decay over time.

There are many such anti-dynamo theorems. The upshot is that although the velocity field may be simple, the magnetic field must be fully three-dimensional. It is thus significant that the observed planetary and stellar magnetic fields exhibit non-axisymmetric components.

3.2 A necessary condition for dynamo action

Note that the existence or non-existence of dynamo action will depend upon the value of $R = UL\sigma\mu$. The magnetic energy is

$$E_m = \int_V \frac{1}{2\mu} |\mathbf{B}|^2 dV, \quad (33)$$

and its rate of change is

$$\begin{aligned} \mu \frac{dE_m}{dt} &= \int_V \mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} dV \\ &= - \int_V \mathbf{B} \cdot \nabla \times \mathbf{E} dV \\ &= - \int_V \mathbf{E} \cdot \nabla \times \mathbf{B} dV \\ &= \int_V \mathbf{u} \cdot (\mathbf{B} \times \nabla \times \mathbf{B}) - \frac{1}{\sigma\mu} |\nabla \times \mathbf{B}|^2 dV. \end{aligned} \quad (34)$$

For a dynamo to exist we need this rate of change of magnetic energy to be non-negative. Now if $U_m = \max_V |\mathbf{u}|$, then we can write

$$\begin{aligned} \int_V \mathbf{u} \cdot (\mathbf{B} \times \nabla \times \mathbf{B}) \, dV &\leq U_m \int_V |\mathbf{B}| |\nabla \times \mathbf{B}| \, dV \\ &\leq U_m \left[\int_V |\mathbf{B}|^2 \, dV \right]^{1/2} \left[\int_V |\nabla \times \mathbf{B}|^2 \, dV \right]^{1/2}, \end{aligned} \quad (35)$$

using the Cauchy-Schwarz inequality. It is known that for a bounded homogeneous conductor V surrounded by vacuum there is a length L_V such that, for any \mathbf{B} in the space of realizable magnetic fields,

$$\int_V |\nabla \times \mathbf{B}|^2 \, dV \geq \frac{1}{L_V^2} \int_{R^3} |\mathbf{B}|^2 \, dV. \quad (36)$$

Thus we have

$$\int_V \mathbf{u} \cdot (\mathbf{B} \times \nabla \times \mathbf{B}) \, dV \leq U_m L_V \int_V |\nabla \times \mathbf{B}|^2 \, dV, \quad (37)$$

so

$$\mu \frac{dE_m}{dt} = \left(U_m L_V - \frac{1}{\sigma \mu} \right) \int_V |\nabla \times \mathbf{B}|^2 \, dV. \quad (38)$$

This is negative if $U_m L_V \sigma \mu < 1$ in which case a dynamo cannot exist. Thus a necessary condition for dynamo action is that $U_m L_V \sigma \mu \geq 1$. In particular dynamo action occurs as a bifurcation from the state of no magnetic field, as the magnetic Reynolds number is increased.

3.3 The Bullard-Gellman computation

There was an early attempt to solve the eigenvalue problem utilizing direct expansion in the models of free decay of a spherical conductor. A steady velocity field was chosen as representative of rotating convection. The truncated system of modal equations produced a positive result, but in 1969 Gibbons and Roberts re-examined this choice of \mathbf{u} and established that it was not in fact a kinematic dynamo. The original result was a numerical artifact.

3.4 Parker's model: The α effect

In 1955, Parker addressed the problem of producing poloidal field from toroidal field, to complement the omega effect. As we have seen, this is impossible if both the velocity field and the magnetic field are axisymmetric, but Parker envisaged averaging of the effects of small up-wellings within the core which could lift and twist poloidal field lines. The local magnetic Reynold's number was taken to be large, so that the magnetic field was 'frozen' into the fluid as it was lifted and twisted (figure 3). Under averaging, the twisted loops produced a mean current, aligned with the local toroidal field line. In this way the poloidal field can be induced from the toroidal field, thus completing the dynamo cycle and maintaining it against dissipation. This mechanism is now known as the ' α effect'. Although Parker's supporting

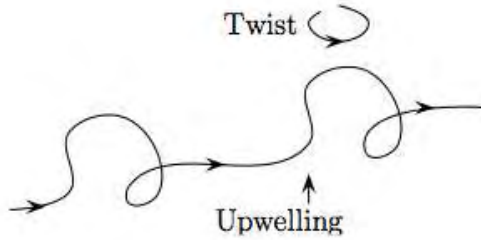


Figure 3: Toroidal field lines are lifted and twisted by small scale non-axisymmetric components of the velocity field. This generates a mean current in the toroidal field direction which induces a poloidal field, and is termed the α effect.

calculations left many gaps, these have since been filled in and the model provides a crucial element of the dynamo cycle, breaking the constraints of Cowling's theorem.

If $\mathbf{u} = U(s)\mathbf{e}_\phi$, the induction equation in this model is written as

$$\mathbf{B} = B(r, z, t)\mathbf{e}_\phi + \nabla \times A(r, z, t)\mathbf{e}_\phi, \quad (39)$$

$$\frac{\partial A}{\partial t} - \frac{1}{R} \left(\nabla^2 - \frac{1}{r^2} \right) A = \alpha B, \quad (40)$$

$$\frac{\partial B}{\partial t} + r \frac{\partial A}{\partial z} \frac{\partial}{\partial r} \left(\frac{U}{r} \right) - \frac{1}{R} \left(\nabla^2 - \frac{1}{r^2} \right) B = 0. \quad (41)$$

The added αB term here is crucial and completes the ' $\alpha - \omega$ dynamo cycle'.

3.5 Early examples of dynamo action

The Herzenberg dynamo (1958), comprises two solid rotating spheres of radius a embedded inside a larger sphere of radius R , with $a \ll R$ and a also small compared to the distance between the two spheres. Each sphere produces an ω effect determined by the local field at that sphere. The resulting induced field increases the local field at the other sphere, and regenerative dynamo action can result. Herzenberg's theoretical analysis of this system relies on the smallness of the embedded spheres and the geometrical decay of the induced field components, leading to a spatial filtering. Such a dynamo was realized experimentally in solid iron by Lowes and Wilkinson (1963).

Backus (1958) gave a proof of dynamo action for a kinematic dynamo in a homogeneous spherical fluid conductor based upon temporal filtering of magnetic decay modes. This involves applying a velocity field $\mathbf{u}_1(\mathbf{x})$ for time $0 < t < T_1$, and then setting $\mathbf{u} = 0$ for time $T_1 < t < T_2$, during which time the magnetic field decays, filtering out modes with the faster decay. A new velocity field $\mathbf{u}_2(\mathbf{x})$ is then applied for $T_2 < t < T_3$ and then $\mathbf{u} = 0$ again. This sequence of applying velocity fields and then setting $\mathbf{u} = 0$ to filter out fast decaying modes is then continued.

If the 'basic mode' (involving the slower decaying poloidal and toroidal components) is \mathbf{D} with $\|\mathbf{D}\| = 1$, an initial magnetic field involving this basic mode with amplitude A , plus

some arbitrary magnetic noise $\mathbf{N}(0)$, with $\|\mathbf{N}(0)\| \leq \epsilon$ for some $\epsilon > 0$, is

$$\mathbf{B}(\mathbf{x}, 0) = A\mathbf{D}(\mathbf{x}) + \mathbf{N}(0). \quad (42)$$

If it can then be shown that at time $t = T_{2N}$ the magnetic field has the form

$$\mathbf{B}(\mathbf{x}, T_{2N}) = A_N\mathbf{D}(\mathbf{x}) + \mathbf{N}(T_{2N}), \quad (43)$$

where $|A_N| \geq |A|$ and $\|\mathbf{N}(T_{2N})\| \leq \epsilon$, then dynamo action occurs. Thus the noise gets no bigger, and the basic field is at least maintained.

3.6 Smoothing Applied to the Kinematic Equation

The smoothing method consists in splitting the dynamo kinematic equation operator

$$\mathcal{L}\mathbf{B} = \frac{\partial}{\partial t}\mathbf{B} - \nabla \times (\mathbf{u} \times \mathbf{B}) - \eta \nabla^2 \mathbf{B} = 0, \quad (44)$$

and the velocity and magnetic field into some rough and smooth components:

$$\begin{aligned} \mathbf{u} &= \mathbf{u}_S + \mathbf{u}_R, \\ \mathbf{B} &= \mathbf{B}_R + \mathbf{B}_S, \\ \mathcal{L} &= \mathcal{L}_R + \mathcal{L}_S, \\ \mathcal{L}_R \mathbf{B} &= -\nabla \times (\mathbf{u}_R \times \mathbf{B}), \\ \mathcal{L}_S \mathbf{B} &= \frac{\partial}{\partial t}\mathbf{B} - \nabla \times (\mathbf{u}_S \times \mathbf{B}) - \eta \nabla^2 \mathbf{B}. \end{aligned} \quad (45)$$

Let P be a projection onto smooth fields ($P^2 = P$), such as

$$P\mathbf{B} = \mathbf{B}_S \quad (46)$$

and assume $P\mathcal{L}_S = \mathcal{L}_S P$. Subtracting $P\mathcal{L}$ from \mathcal{L} , one gets

$$-(\mathcal{L}_R - P\mathcal{L}_R)\mathbf{B} = (\mathcal{L}_S - P\mathcal{L}_S)\mathbf{B} = \mathcal{L}_S\mathbf{B} - \mathcal{L}_S\mathbf{B}_S = \mathcal{L}_S\mathbf{B}_R \quad (47)$$

The smooth projection of $(\mathcal{L}_R - P\mathcal{L}_R)\mathbf{B}$ is

$$P(\mathcal{L}_R - P\mathcal{L}_R)\mathbf{B} = P\mathcal{L}_R - P^2\mathcal{L}_R\mathbf{B} = P\mathcal{L}_R - P\mathcal{L}_R\mathbf{B} = 0 \quad (48)$$

so $(\mathcal{L}_R - P\mathcal{L}_R)\mathbf{B}$ is a rough field. Assuming that \mathcal{L}_S can be inverted on rough fields, we have

$$\mathbf{B}_R = -\mathcal{L}_S^{-1}(\mathcal{L}_R - P\mathcal{L}_R)\mathbf{B} \equiv M\mathbf{B} \quad (49)$$

So, one has

$$\mathbf{B} - \mathbf{B}_S = \mathbf{B}_R = M\mathbf{B} \quad (50)$$

or, if $I - M$ is invertible on smooth fields, we obtain \mathbf{B} in terms of its smooth component,

$$\mathbf{B} = (I - M)^{-1}\mathbf{B}_S \quad (51)$$

Thus, one can write

$$\begin{aligned} P\mathcal{L}\mathbf{B} &= P(\mathcal{L}_S + \mathcal{L}_R)\mathbf{B} = \mathcal{L}_S\mathbf{B}_S + P\mathcal{L}_R(I - M)^{-1}\mathbf{B}_S \\ &= [\mathcal{L}_S + P\mathcal{L}_R(I - M)^{-1}]\mathbf{B}_S = 0 \end{aligned} \quad (52)$$

and it can be shown by substitution that if \mathbf{B}_S satisfies

$$[\mathcal{L}_S + P\mathcal{L}_R(I - M)^{-1}]\mathbf{B}_S = 0 \quad (53)$$

then $\mathbf{B} = (I - M)^{-1}\mathbf{B}_S$ solves $\mathcal{L}\mathbf{B} = 0$. The smoothing method can be used to treat analytically the creation of the α -effect.

3.6.1 First-Order Smoothing

Assuming that $\mathcal{L}_R\mathbf{B}_S$ is a rough field such as $P\mathcal{L}_R\mathbf{B}_S = 0$, one can make the following approximations

$$0 = [\mathcal{L}_S + P\mathcal{L}_R(I - M)^{-1}]\mathbf{B}_S \approx [\mathcal{L}_S + P\mathcal{L}_R(I + M)]\mathbf{B}_S = [\mathcal{L}_S + P\mathcal{L}_R M]\mathbf{B}_S. \quad (54)$$

Here, we have

$$\mathcal{L}_S\mathbf{B}_S = \frac{\partial}{\partial t}\mathbf{B}_S - \nabla \times (\mathbf{u}_S \times \mathbf{B}_S) - \frac{1}{R}\nabla^2\mathbf{B}_S \quad (55)$$

and

$$P\mathcal{L}_R M\mathbf{B}_S = -P[\mathcal{L}_R\mathcal{L}_S^{-1}\mathcal{L}_R]\mathbf{B}_S \approx P\mathcal{L}_R\mathbf{B}_R \quad (56)$$

with \mathbf{B}_R a linear function of \mathbf{B}_S . This is the basis of mean field electrodynamics developed by Krause, Radler and others in the 60's.

For example, set $\mathbf{u}_S = 0$ and let \mathbf{u}_R be a real steady solenoidal field of the form

$$\mathbf{u}_R = \sum_{k \in K} \mathbf{a}_k e^{i\mathbf{k} \cdot \mathbf{x}} \quad (57)$$

where K is all 3-vectors excluding the zero vector. If \mathbf{B}_S is large scale relative to \mathbf{u}_R , then

$$\nabla \times (\mathbf{u}_R \times \mathbf{B}_S) = -\frac{1}{R}\nabla^2\mathbf{B}_R \quad (58)$$

implies

$$\mathbf{B}_R = Ri \sum_{k \in K} k^{-2} \mathbf{k} \cdot \mathbf{B}_S \mathbf{a}_k e^{i\mathbf{k} \cdot \mathbf{x}}, \quad (59)$$

and

$$P[\mathbf{u}_R \times \mathbf{B}_R] \approx Ri \sum_{k \in K} k^{-2} (\mathbf{a}_k^* \times \mathbf{a}_k) \mathbf{B}_S \cdot \mathbf{k} \equiv \alpha \cdot \mathbf{B}_S. \quad (60)$$

It can be shown, by using the substitution $\mathbf{a}_k = \mathbf{k} \times \mathbf{b}_k$, that α is real and symmetric. For isotropic \mathbf{u}_R , of this form we may assume $\alpha_{ij} = \alpha \delta_{ij}$ where

$$\alpha = \frac{1}{3} Ri \sum_{\mathbf{k} \in R} k^{-2} (\mathbf{a}_k^* \times \mathbf{a}_k) \cdot \mathbf{k}. \quad (61)$$

If all \mathbf{k} have unit length we see that

$$\alpha = \frac{R}{3} \overline{\mathbf{u}_R \cdot (\nabla \times \mathbf{u}_R)}, \quad (62)$$

involving a spatial average of the helicity density $\mathbf{u}_R \cdot (\nabla \times \mathbf{u}_R)$ of the velocity field. A field satisfying all the above conditions is the Beltrami field (in which \mathbf{u} is parallel to $\nabla \times \mathbf{u}$)

$$\mathbf{u}_R = (\sin z + \cos y, \sin x + \cos z, \sin y + \cos x), \quad (63)$$

with $\alpha = R/3$. The role of helicity is reminiscent of Parker's 'upwellings', which involved a rising, twisting flow.

3.6.2 Spatially periodic kinetic dynamos

This is a setting where complete analysis of the smoothing method may be carried out explicitly, with or without time dependence. Considering here only steady fields, the real velocity field has the form

$$\mathbf{u} = \mathbf{u}_R = \sum_{k \in K} \mathbf{a}_k e^{i\mathbf{k} \cdot \mathbf{x}} \quad (64)$$

The admissible magnetic fields have the complex form

$$\mathbf{B} = e^{i\mathbf{n} \cdot \mathbf{x}} \sum_{k \in K + (0,0,0)} \mathbf{B}_k e^{i\mathbf{k} \cdot \mathbf{x}}, |\mathbf{n}| < 1 \quad (65)$$

so $\mathbf{B}_0 = \mathbf{B}_S$ (the components for $k \neq 0$ are all rough). Kinematic dynamo action can be defined by growth of field if n is sufficiently small. This is the palindromic 'so many dynamos' theorem (G. O. Roberts): almost all such \mathbf{u} are kinematic dynamos.

3.7 Braginskii's theory for nearly axisymmetric dynamos (1964)

Braginskii used a smoothing technique for nearly axisymmetric flows to derive the α effect. Here, the 'smooth' parts of the variables are the axisymmetric components, and the rough parts are the non-axisymmetric components. The smoothing operation is therefore an angle average over ϕ in cylindrical polar coordinates. Motivated by the fact that the magnetic Reynolds number R is large for the Earth and that the field is dominated by a dipole aligned with the rotation axis, Braginskii considers expanding the fields in powers of $\epsilon = R^{-1/2}$;

$$\mathbf{u} = U(r, z) \mathbf{e}_\phi + \epsilon \mathbf{u}' + \epsilon^2 \mathbf{u}_P, \quad (66)$$

$$\mathbf{B} = B(r, z, t) \mathbf{e}_\phi + \epsilon \mathbf{B}' + \epsilon^2 \mathbf{B}_P, \quad (67)$$

where $\langle \mathbf{u}' \rangle = \langle \mathbf{B}' \rangle = 0$ where $\langle \cdot \rangle$ means an angle average. The relevant timescale is shorter, so writing $\tau = \epsilon^2 t$, the order ϵ terms in the induction equation give

$$\frac{\partial B}{\partial \tau} + r \mathbf{u}_P \cdot \nabla \left(\frac{B}{r} \right) = r \mathbf{B}_P \cdot \nabla \left(\frac{U}{r} \right) + (\nabla \times \mathcal{E})_\phi + \left(\nabla^2 - \frac{1}{r^2} \right) B = 0, \quad (68)$$

where $\mathcal{E} = \langle \mathbf{u}' \times \mathbf{B}' \rangle$. If $\mathbf{B}_P = \nabla \times A(r, z, t) \mathbf{e}_\phi$, we also find

$$\frac{\partial A}{\partial \tau} + r \mathbf{u}_P \cdot \nabla \left(\frac{A}{r} \right) = \mathcal{E}_\phi + \left(\nabla^2 - \frac{1}{r^2} \right) A. \quad (69)$$

Braginskii found that the contributions from \mathcal{E} can be mostly absorbed into ‘effective’ variables, denoted with a superscript e :

$$\frac{\partial B}{\partial \tau} + r \mathbf{u}_P^e \cdot \nabla \left(\frac{B}{r} \right) = r \mathbf{B}_P \cdot \nabla \left(\frac{U}{r} \right) + \left(\nabla^2 - \frac{1}{r^2} \right) B = 0, \quad (70)$$

$$\frac{\partial A^e}{\partial \tau} + r \mathbf{u}_P^e \cdot \nabla \left(\frac{A^e}{r} \right) = \alpha B + \left(\nabla^2 - \frac{1}{r^2} \right) A^e. \quad (71)$$

The term which cannot be absorbed gives rise to the α term, and in effective variables these equations are identical to the Parker model, the toroidal field exhibiting the α effect.

3.8 Soward’s Lagrangian analysis of Braginskii’s problem

The same situation of a nearly axisymmetric flow was analysed in a slightly different way by Soward (1972). The key idea of this method is to make use of an invariance of the induction equation for the diffusionless case $R = \infty$;

$$\mathbf{B}_t + \mathbf{u} \cdot \nabla \mathbf{B} = \mathbf{B} \cdot \nabla \mathbf{u}. \quad (72)$$

Consider two (prescribed) velocity fields $\mathbf{u}(\mathbf{x}, t)$ and $\tilde{\mathbf{u}}(\tilde{\mathbf{x}}, t)$, and the Lagrangian variables $\mathbf{x}(\mathbf{a}, t)$ and $\tilde{\mathbf{x}}(\mathbf{a}, t)$ for these flows defined by

$$\left. \frac{d\mathbf{x}}{dt} \right|_{\mathbf{a}} = \mathbf{u}(\mathbf{x}, t), \quad \left. \frac{d\tilde{\mathbf{x}}}{dt} \right|_{\mathbf{a}} = \tilde{\mathbf{u}}(\tilde{\mathbf{x}}, t). \quad (73)$$

Then

$$\tilde{u}_i = \left. \frac{\partial \tilde{x}_i}{\partial t} \right|_{\mathbf{x}} + \frac{\partial \tilde{x}_i}{\partial x_j} u_j. \quad (74)$$

The solutions to the induction equation for the two flows are

$$B_i(\mathbf{x}, t) = B_j(\mathbf{a}, 0) \frac{\partial x_i}{\partial a_j}, \quad \tilde{B}_i(\tilde{\mathbf{x}}, t) = \tilde{B}_j(\mathbf{a}, 0) \frac{\partial \tilde{x}_i}{\partial a_j}. \quad (75)$$

Using the chain rule

$$\frac{\partial \tilde{x}_i}{\partial a_j} = \frac{\partial \tilde{x}_i}{\partial x_k} \frac{\partial x_k}{\partial a_j} \quad (76)$$

we obtain

$$\tilde{B}_i = \frac{\partial \tilde{x}_i}{\partial x_k} B_k \quad (77)$$

Now the idea is to consider one of the velocity fields as including the ‘rough’ components while the other is the smooth velocity field without these rough components. For the Braginskii problem this means taking the tilde variables to include non-axisymmetric components of order ϵ , while the \mathbf{x} variables do not include these. The \mathbf{x} variables will therefore not be

physical. We seek a near identity transformation $\mathbf{x}(\tilde{\mathbf{x}}, t) = \tilde{\mathbf{x}} + O(\epsilon)$ of R^3 which eliminates the $O(\epsilon)$ in the velocity, so that

$$\mathbf{u}(\mathbf{x}, t) = U(r, x)(1 + O(\epsilon^2))\mathbf{e}_\phi + \epsilon^2 \mathbf{u}_p^e(\mathbf{x}) + \epsilon^2 \mathbf{u}''(\mathbf{x}, t) + o(\epsilon^2). \quad (78)$$

The equations in \mathbf{x} space can be angle averaged, and this results in Braginskii's result (70) and (71) with $A(\mathbf{x}, \tau)$ in place of $A^e(\tilde{\mathbf{x}}, \tau)$. The α effect is missing - it comes from the non-invariant diffusive term. The effective variable A^e occurs when we transform the unphysical \mathbf{x} variables back into the real physical space $\tilde{\mathbf{x}}$.

The key point is that, by including the non-invariant diffusive term in this analysis, one obtains the α effect term in (71), along with terms which can be absorbed into effective variables.

4 Some examples of analytically accessible dynamos

There are some simple velocity fields for which dynamo action is known to occur. The Ponomarenko dynamo is a flow in a cylinder in which the fluid travels axially while rotating,

$$\mathbf{u} = U(r)\mathbf{e}_z + r\Omega(r)\mathbf{e}_\phi \quad (79)$$

The magnetic field is periodic in z ,

$$\mathbf{B} = b(r)e^{\lambda t + im\phi + ikz}. \quad (80)$$

This can be analysed for large R . The dynamo mechanism here involves the Ω effect which maintains the toroidal field from the poloidal field, and diffusion which maintains the poloidal field from the toroidal field.

Another simple dynamo is the Roberts cell, which like the Beltrami flow is periodic in two dimensions

$$\mathbf{u} = (\cos y, \sin x, \cos x + \sin y). \quad (81)$$

By transforming the variables $x \rightarrow x - y$, $y \rightarrow \pi/2 + x + y$, the Roberts cell takes the form

$$\mathbf{u} = (\cos y \sin x, -\cos x \sin y, \sqrt{2} \sin x \sin y) \quad (82)$$

An illustration of this flow and the generation of magnetic field is shown in figure 4.

5 MHD dynamos

So far we have only talked about *kinetic* dynamo theory, when the velocity field is somehow prescribed and unaffected by the magnetic field. In magnetohydrodynamics the velocity and magnetic fields both interact with each other and some physical mechanism (such as convection) which drives the fluid velocity must be included. The momentum equation for the fluid in a rotating frame with angular velocity Ω includes the Coriolis force and the Lorentz force,

$$\rho \frac{D\mathbf{u}}{Dt} + \nabla p + 2\rho\Omega \times \mathbf{u} + \frac{1}{\mu}\mathbf{B} \times (\nabla \times \mathbf{B}) - \rho\nu\nabla^2\mathbf{u} = c\mathbf{g}. \quad (83)$$

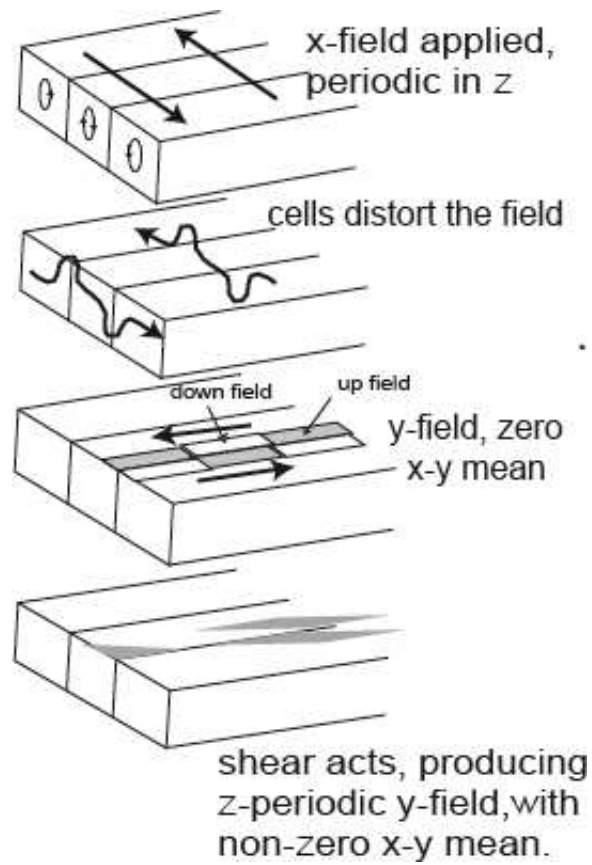


Figure 4: The Roberts cell with a z periodic field applied. The flow causes the production of z periodic y -field from x -field, and vice versa.

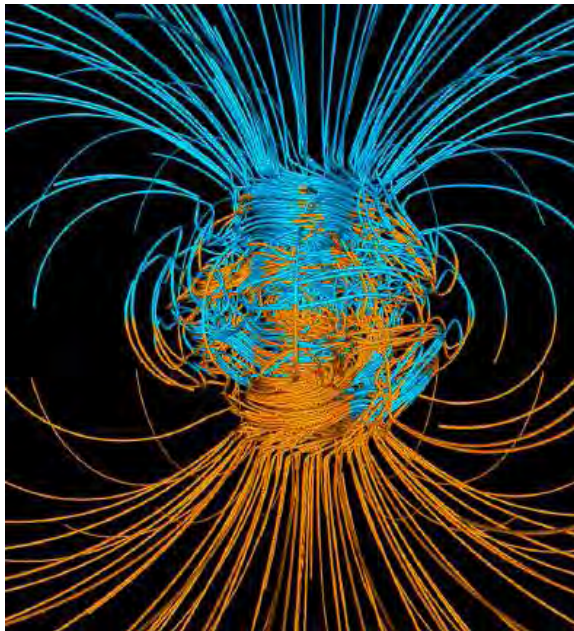


Figure 5: Magnetic field from a numerical MHD simulation of the Earth’s core by Glatzmaier and Roberts.

The motion here is assumed to be driven by convection of some scalar c (temperature or composition, for instance),

$$\frac{Dc}{Dt} - D_c \nabla^2 c = Q. \quad (84)$$

If the convective forces are absent, and the velocity is axisymmetric, but the alpha-effect is included as in Braginskii’s model, and the normal component of \mathbf{u} vanishes at the core boundary, it is found that such an axisymmetric dynamo cannot be sustained in MHD. The microscale driving needed to produce the small scale fields associated with an alpha effect can not sustain the dynamo. The Earth’s dynamo is now believed to be a result of larger scale driving as a result of thermal and compositional convection associated with the formation of the inner core.

Figure 5 shows a numerical simulation by Glatzmaier and Roberts, of a convective model of the Earth’s core. This simulation exhibits the observed full reversals of the magnetic field on timescales of geophysical relevance. The role of the inner core is established as a reservoir of magnetic field during these reversals.

References

- [1] E. C. BULLARD, *The Origin of the Earth’s Magnetic Field*, Observatory, 70 (1950), pp. 139–143.
- [2] T. G. COWLING, *The Magnetic Field of sunspots*, Monthly Notices R. Astron. Soc., 94 (1933), pp. 39–48.

- [3] T. G. COWLING, *Magnetohydrodynamics*, Interscience Publishers, London, 1957.
- [4] W. M. ELSASSER, *Hydromagnetism. I. A Review*, Am. J. Phys., 23 (1955), pp. 590–608.
- [5] M. GHIL AND S. CHILDRESS, *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics*, Springer-Verlag, New York, 1987.
- [6] H. K. MOFFATT, *Magnetic Field Generation in Electrically Conducting Fluids*, Cambridge University Press, 1978.
- [7] E. N. PARKER, *Hydrodynamic Dynamo models*, Astrophys. J., 122 (1955), pp. 293–314.
- [8] E. N. PARKER, *Cosmical Magnetic Fields*, Calrendon Press, Oxford, 1979.
- [9] G. O. ROBERTS, *Spatially periodic dynamos*, Phil. Trans. R. Soc. A, 266 (1970), pp. 535–558.
- [10] P. H. ROBERTS, *An Introduction to Magnetohydrodynamics*, Americal Elsevier Publishing Company, New York, 1967.
- [11] P. H. ROBERTS, *Dynamo Theory*, vol. 2, American Mathematical Society, Providence, RI, 1971, pp. 129–206.

Convection, Stability and Turbulence

Lecturer: Charlie Doering

Notes written by: Chris Cawthorn & Toby Wood

18 June 2008

1 Introduction

The three basic forms of heat transport are:

- Radiation
- Conduction
- Convection

In many physical systems, all of these processes may play a role (see Figures 1–2 below). Convection is buoyancy-driven flow due to density variations in a gravitational field. It is particularly important as it enhances the vertical heat transport. For example, without convection, it would take many hours to heat a saucepan of water to boiling point, simply because heat diffusion in water is so weak ($\kappa \approx 1.5 \times 10^{-3} \text{cm}^2 \text{s}^{-1}$).

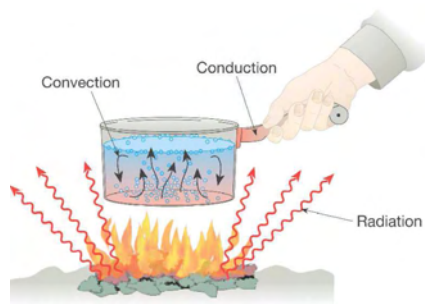


Figure 1: Heat transfer in a saucepan.

2 Rayleigh-Bénard convection

In 1900, Henri Bénard [2] performed an experiment concerning convection cells in a thin liquid layer. He observed spontaneous pattern formation - the convection was seen to organise itself into hexagonal cells throughout the entire domain, as shown in Figure 3.



Figure 2: Heat transfer in the Earth and Sun.



Figure 3: Convection cells in Bénard's experiment and on the surface of the Sun.

Rayleigh [28] was the first to undertake a mathematical theory of fluid convection. His analysis described the formation of convection rolls for fluid confined between parallel plates held at different temperatures. The width of the rolls at convective onset is proportional to the depth of the fluid layer, but also depends on details of the boundary conditions.

However, these convection rolls are a different physical phenomenon to the convection cells observed by Bénard. Such convection cells are now known to be associated with Marangoni flows, where variations of surface tension with temperature play a crucial role. In fact, there are several regimes of convection. As the temperature difference is increased, spiral defect chaos may be observed, before finally giving way to fully turbulent convection, with the associated breakdown of spacial-temporal correlations (see Figure 4).

2.1 A mathematical theory of convection

The theory first developed by Rayleigh is now known as Rayleigh–Bénard convection. Provided that the temperature variations within the fluid layer remain small, we may adopt the Boussinesq approximation, which neglects compressibility except for the presence of a buoyancy term in the momentum equation. Furthermore, it may be shown that viscous heating is often negligible compared to thermal driving from the boundary conditions. The

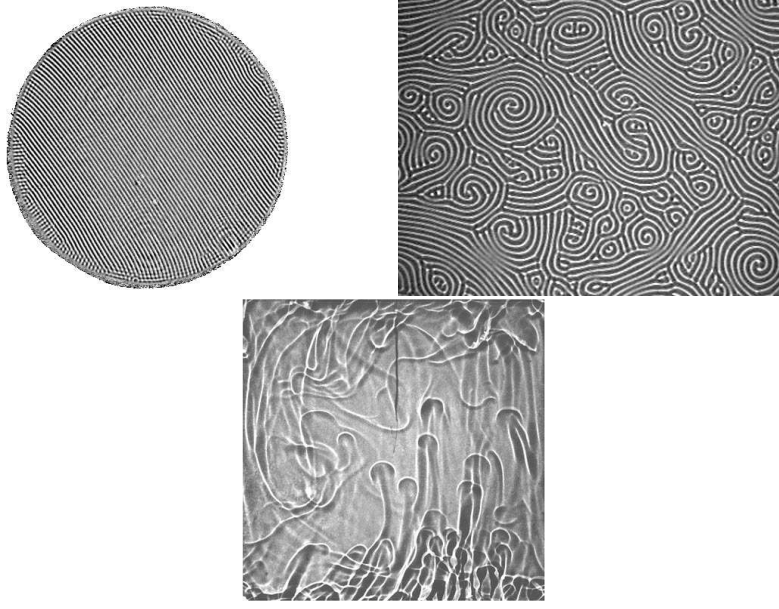


Figure 4: Convection rolls give way to complex spatio-temporal dynamics, and eventually to fully-developed turbulent convection. (The last figure is a side view.)

equations of mass, momentum and heat conservation are

$$0 = \nabla \cdot \mathbf{u}, \quad (1)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{1}{\rho_0} \nabla(p - p_0) = \nu \nabla^2 \mathbf{u} + g \alpha \hat{\mathbf{k}}(T - T_0), \quad (2)$$

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = \kappa \nabla^2 T. \quad (3)$$

The last equation arises from the energy equation

$$\frac{\partial q}{\partial t} + \nabla \cdot \mathbf{J} = 0 \quad (4)$$

where $q = cT$ is the thermal energy density and $\mathbf{J} = c(\mathbf{u}T - \kappa \nabla T)$ is the heat flux. In equation (2) ν is kinematic viscosity, g is acceleration due to gravity, c is specific heat capacity and α is the thermal expansion coefficient, which also appears in the equation of state

$$\frac{\rho - \rho_0}{\rho_0} = \alpha(T - T_0). \quad (5)$$

Rayleigh considered a domain confined between two impermeable flat horizontal plates, each held at a different temperature. The distance between the plates is h , and the problem is assumed to be periodic in the two horizontal directions, x and y (see Figure 5). We are primarily concerned with the mean vertical heat flux, namely

$$\langle J_z \rangle = \left\langle wcT - c\kappa \frac{\partial T}{\partial z} \right\rangle \quad (6)$$

$$= \frac{c\kappa}{h}(T_{\text{hot}} - T_{\text{cold}}) + c \langle wT \rangle \quad (7)$$

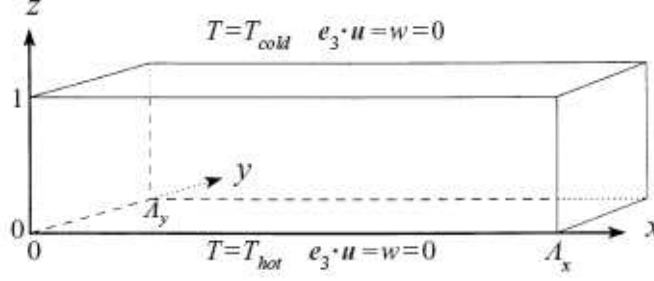


Figure 5: The mathematical domain.

where $\langle \cdot \rangle$ denotes a space-time average.

From this point onwards, we shall work with non-dimensionalised equations. The characteristic length scale of the system is the height h , and the characteristic time scale is the thermal diffusion time over this distance, h^2/κ . We therefore formally define the following dimensionless variables:

$$\tilde{\mathbf{x}} = \mathbf{x}/h \quad (8)$$

$$\tilde{t} = \kappa t/h^2 \quad (9)$$

$$\tilde{\mathbf{u}} = \mathbf{u}h/\kappa \quad (10)$$

$$\tilde{T} = \frac{T - T_0}{T_{\text{hot}} - T_{\text{cold}}} \quad (11)$$

$$\tilde{p} = \frac{p - p_0}{\rho_0} \frac{h^2}{\kappa^2} \quad (12)$$

If we further define the dimensionless parameters

- Rayleigh number, $\text{Ra} = \frac{g\alpha(T_{\text{hot}} - T_{\text{cold}})h^3}{\nu\kappa}$
- Prandtl number, $\text{Pr} = \nu/\kappa$

then the governing equations become (after dropping the ‘ \sim ’s)

$$0 = \nabla \cdot \mathbf{u}, \quad (13)$$

$$\frac{1}{\text{Pr}} \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) + \nabla p = \nabla^2 \mathbf{u} + \text{Ra} \, \hat{\mathbf{k}} T, \quad (14)$$

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = \nabla^2 T \quad (15)$$

The boundary conditions on T are now simply $T = 1$ on $z = 0$ and $T = 0$ on $z = 1$. Finally, we define the Nusselt number, Nu , which measures the enhancement of heat transport due to convection, relative to conductive heat transfer:

$$\text{Nu} = \langle J_z \rangle / J_{\text{cond}} \quad (16)$$

We wish to determine how Nu varies with the parameters Ra and Pr ¹. From the equations of motion, it can be shown (see Appendix A) that

$$\text{Nu} = 1 + \langle wT \rangle \quad (17)$$

$$= \langle |\nabla T|^2 \rangle \quad (18)$$

$$= 1 + \frac{1}{\text{Ra}} \langle |\nabla \mathbf{u}|^2 \rangle \geq 1, \quad (19)$$

the last expression showing that convection always *increases* the rate of heat transfer. If we chose to represent the convective heat flux in terms of an effective “eddy diffusion” κ_{eddy} , so that $\langle J_z \rangle = c\kappa_{\text{eddy}}(T_{\text{hot}} - T_{\text{cold}})/h$, we would then find that $\kappa_{\text{eddy}} = \text{Nu} \kappa$. In the case of no convection we have the “conduction solution”, $\mathbf{u} = 0$, $T = \tau(z) = 1 - z$, with $\text{Nu} = 1$.

2.2 Linear instability

As we increase the temperature difference between the two plates (i.e. increase the Rayleigh number) we expect the conduction solution to become unstable. We therefore analyse the linear stability of the system by defining a perturbation θ such that $T = \tau(z) + \theta(\mathbf{x}, t)$. We now linearise our equations in order to obtain a single linear equation for θ :

$$(-\frac{1}{\text{Pr}}\partial_t + \nabla^2)(-\partial_t + \nabla^2)\nabla^2\theta - \text{Ra}(\nabla^2 - \partial_z^2)\theta = 0. \quad (20)$$

We look for solutions of the form $\theta \propto e^{-\lambda t + ik_x x + ik_y y}$, which yields the 1D eigenfunction problem

$$(\frac{1}{\text{Pr}}\lambda + \partial_z^2 - k^2)(\lambda + \partial_z^2 - k^2)(\partial_z^2 - k^2)\theta + \text{Ra} k^2\theta = 0, \quad (21)$$

where $k^2 = k_x^2 + k_y^2$. On both plates we have $\theta = 0$ and $\partial_z^2\theta = 0$ so that $w = 0$. If we further assume, for simplicity, that we have no-stress boundary conditions on the two plates, then the solutions of (21) are $\theta \propto \sin(n\pi z)$ for $n \in \mathbb{Z}$. We find that $\lambda \in \mathbb{R} \forall k, n$, so the condition for marginal stability is $\lambda = 0$. The marginally stable modes have

$$\text{Ra} = \frac{(k^2 + n^2\pi^2)^3}{k^2}, \quad (22)$$

and we find the critical Rayleigh number Ra_c (above which the system is unstable to disturbances of some wavelength) by minimising this expression over k and n . Therefore, regardless of the choice of Pr , we find $\text{Ra}_c = \frac{27}{4}\pi^4 \approx 657$, at which point convective rolls appear with width $\pi/k_c = \sqrt{2}h$. If we instead choose no-slip boundary conditions on the two plates, then the eigenfunctions become more complicated, and the critical Rayleigh number increases to $\text{Ra}_c \approx 1708$.

2.3 Nonlinear stability

Linear stability does not rule out the possibility that the system might become nonlinearly unstable at Rayleigh numbers smaller than Ra_c . In order to determine sufficient conditions for *stability*, we can perform an “energy analysis”, retaining all terms in the equations. We

¹In a domain of finite horizontal extent, the cross-sectional area would also enter as a parameter.

again suppose that $T = \tau(z) + \theta(\mathbf{x}, t)$, where $\tau(z) = 1 - z$, and consider perturbations that are periodic in x and y . By taking the scalar product of \mathbf{u} with (14) and applying (13) we find

$$\frac{\partial}{\partial t} \left(\frac{u^2}{2\text{Pr}} \right) + \nabla \cdot \left(\frac{u^2}{2\text{Pr}} \mathbf{u} + p\mathbf{u} \right) = \nabla^2 \left(\frac{1}{2} u^2 \right) - |\nabla \mathbf{u}|^2 + \text{Ra } Tw. \quad (23)$$

Integrating over the domain and applying the boundary conditions, we find

$$\frac{1}{\text{Pr Ra}} \frac{d}{dt} \int_V \frac{1}{2} u^2 dV = -\frac{1}{\text{Ra}} \int_V |\nabla \mathbf{u}|^2 dV + \int_V \theta w dV. \quad (24)$$

Similarly, we can multiply (15) by $\theta = T - \tau(z)$ to obtain

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \theta^2 \right) + \nabla \cdot \left(\frac{1}{2} \theta^2 \mathbf{u} \right) = -(\tau' w - \tau'') \theta + \nabla^2 \left(\frac{1}{2} \theta^2 \right) - |\nabla \theta|^2, \quad (25)$$

which becomes, after integration,

$$\frac{d}{dt} \int_V \frac{1}{2} \theta^2 dV = - \int_V |\nabla \theta|^2 dV - \int_V (\tau' w - \tau'') \theta dV. \quad (26)$$

Since $\tau(z) = 1 - z$, the sum of equations (24) and (26) is

$$\frac{dE}{dt} = -Q \quad (27)$$

$$= -(Q/E)E \leq - \left(\min_{\theta, \mathbf{u}} [Q/E] \right) E \quad (28)$$

where

$$E\{\theta, \mathbf{u}\} = \int_V \frac{1}{2} \left[\frac{u^2}{\text{Pr Ra}} + \theta^2 \right] dV \quad (29)$$

and

$$Q\{\theta, \mathbf{u}\} = \int_V \left[\frac{|\nabla \mathbf{u}|^2}{\text{Ra}} + |\nabla \theta|^2 - 2\theta w \right] dV. \quad (30)$$

If we can show that $\min_{\theta, \mathbf{u}} [Q/E]$ is strictly greater than zero (for given Ra), then equation (28) implies that E tends exponentially to zero, so the system is *absolutely stable*. Since both E and Q are quadratic forms, it suffices to minimise over θ and \mathbf{u} such that $E\{\theta, \mathbf{u}\} = 1$. We therefore seek to extremise the functional

$$\mathcal{F}\{\theta, \mathbf{u}\} = Q\{\theta, \mathbf{u}\} - 2\lambda E\{\theta, \mathbf{u}\} - \int_V \frac{2}{\text{Ra}} p(\mathbf{x}) \nabla \cdot \mathbf{u} dV \quad (31)$$

where λ and $p(\mathbf{x})$ are, respectively, global and pointwise Lagrange multipliers, designed to ensure that $E = 1$ and $\nabla \cdot \mathbf{u} = 0$ throughout the domain. The Euler-Lagrange equations for this system are

$$-\frac{\lambda}{\text{Pr}} \mathbf{u} + \nabla p = \nabla^2 \mathbf{u} + \text{Ra } \theta \hat{\mathbf{k}} \quad (32)$$

$$-\lambda \theta - w = \nabla^2 \theta \quad (33)$$

From these, it can be shown that $Q\{\theta, \mathbf{u}\} = 2\lambda$, so it just remains to show that λ is positive. In fact, we can combine (32–33) into the single equation

$$\left(\frac{1}{\text{Pr}}\lambda + \nabla^2\right)(\lambda + \nabla^2)\nabla^2\theta - \text{Ra}(\nabla^2 - \partial_z^2)\theta = 0. \quad (34)$$

Comparing this to equation (20), we see that the nonlinear stability criterion is *identical* to the linear stability criterion.

In principle, we could repeat this analysis for a different functional form of $\tau(z)$ in equation (26). We shall see in §4.1 that this approach allows us to derive upper bounds for the Nusselt number Nu when $\text{Ra} > \text{Ra}_c$.

3 Heat transport at high Ra

In the decades following Rayleigh’s discovery, most studies of convection focused on the weakly nonlinear regime at Rayleigh numbers just above critical. The first author to study fully turbulent convection was Malkus [18]. Later, Kraichnan [16] applied mixing length theory to predict the behaviour of the Nusselt number in the limit of large Rayleigh number.

A more extensive theory was proposed by Grossmann & Lohse [12], which led to a complicated picture with eight separate parameter regimes, depending on the values of Ra and Pr (see Figure 6). The “ultimate” scaling obtained in the limit of large Rayleigh number

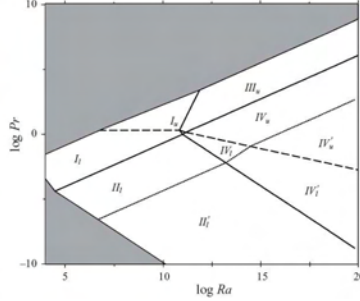


Figure 6: The various parameter regimes predicted by [12].

was predicted by both Kraichnan and Grossmann & Lohse to be $\text{Nu} \sim (\text{Pr Ra})^{1/2}$.

3.1 The “ultimate” scaling: $\text{Nu} \propto (\text{Pr Ra})^{1/2}$

In fact, the “ultimate” scaling law can be derived from the so-called “free-fall” argument of Spiegel [31]. If the diffusivities are sufficiently small, then the rate of heat flux may be determined by the bulk dynamics. The vertical velocity scale is therefore set by balancing inertia and buoyancy acceleration:

$$w^2/h \sim g\alpha\delta T. \quad (35)$$

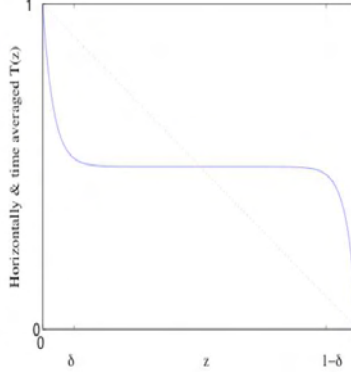


Figure 7: A sketch of the mean temperature in a turbulently-convecting Rayleigh-Bénard cell.

Since the thermal energy density is $c\delta T$, the effective vertical heat flux is then $J_{conv} \sim c\delta T w$, which is independent of ν and κ . Thus

$$\text{Nu} \sim \frac{J_{conv}}{J_{cond}} \quad (36)$$

$$\sim \frac{c\Delta T [g\alpha h \Delta T]^{1/2}}{c\kappa \Delta T / h} \quad (37)$$

$$\sim (\text{Pr Ra})^{1/2}. \quad (38)$$

3.2 The “classical” scaling: $\text{Nu} \propto \text{Ra}^{1/3}$

An argument in favour of the scaling $\text{Nu} \propto \text{Ra}^{1/3}$ was originally proposed by Malkus [19], and quantitatively articulated by Howard [13]. Howard’s argument, which appeals to a plume-like, time-dependent solution of the heat equation, is omitted here in favour of a simpler, though less precise, approach.

In a statistically steady state, the mean temperature in a turbulently-convecting cell should be uniform in the well-mixed interior, with most of the variation occurring in thin boundary layers of thickness δ at the top and bottom of the cell, as illustrated in Figure 7. In dimensionless variables, the heat flux (hence the Nusselt number Nu) away from the lower boundary will be proportional to the gradient of T there, namely δ^{-1} . In order to determine the boundary layer thickness, we assume *marginal stability*. That is, we assume that the Rayleigh number based on conditions in the boundary layer is (almost) exactly critical. Thus

$$\text{Ra}_c \approx \text{Ra}_\delta = \frac{\alpha g (\delta h)^3 \Delta T}{2\nu\kappa} = \frac{1}{2} \delta^3 \text{Ra} \quad (39)$$

$$\Rightarrow \delta \approx (2\text{Ra}_c)^{1/3} \text{Ra}^{-1/3} \quad \Rightarrow \text{Nu} \propto \text{Ra}^{1/3}. \quad (40)$$

An interesting aside is that the prediction made here for the prefactor is fairly accurate, lending further weight to this simple argument.

3.3 Experimental and numerical validation of the scaling laws

Recent years have seen several experimental works attempt to verify the various proposed scalings for heat transport at high Rayleigh number. In experiments with gaseous helium, Chavanne *et al.* [4] claimed to have observed a distinct high-Ra (in their case, $Ra > 10^{11}$) scaling for Nu, perhaps matching the predictions of Kraichnan. Their results were later contradicted by Glazier *et al.* [11], who observed no deviation from another scaling, $Ra^{2/7}$, in their experiments with liquid mercury, though the experimental Prandtl number was much smaller than in [4]. Experimental evidence for a $Ra^{0.301} \approx Ra^{3/10}$ scaling (or perhaps a $(Ra \log Ra)^{3/10}$ scaling) was obtained by Niemela *et al.* [21]. High-precision experiments carried out by Nikolaenko *et al.* [22] suggest that the $Ra^{1/3}$ scaling can be observed for convecting cells with large aspect ratio. While the debate over the correct scaling is by no means settled, the three-dimensional simulations of Amati *et al.* [1], reproduce the classical $Ra^{1/3}$ scaling.

The somewhat ephemeral nature of the high-Ra scaling may perhaps be indicative of the fact that boundaries play a non-trivial role in determining turbulent heat transport, accentuating differences in experimental apparatus. A sufficiently rough boundary may disrupt any boundary layers, changing the behaviour of the system. Experiments in a rough-walled container by Roche *et al.* [29] produced evidence for the $Ra^{1/2}$ scaling, a result that was reproduced by the numerical simulations of Stringano *et al.* [32]. Notably, their simulation yielded *different* results when their jagged bottom boundary was replaced by a smooth, flat boundary.

4 Analytical upper bounds on Nu

With so many experiments giving different results, one would like to investigate the possibility of calculating mathematical upper bounds for the Nusselt number Nu. This can be done by using the integral identities for Nu (17-19) and applying standard inequalities from calculus in order to obtain bounds. One method of carrying out such a calculation is the so-called *background method*, which is outlined below.

4.1 The background method

We begin by allowing the temperature field T to take the form

$$T(\mathbf{x}, t) = \tau(z) + \theta(\mathbf{x}, t) \quad (41)$$

We are entirely free to choose $\tau(z)$, as long as it satisfies the inhomogeneous boundary conditions. Unlike in the energy analysis of stability described earlier in §2.3, it need not satisfy the steady-state conduction equation (with $\mathbf{u} = 0$). Unlike a linear stability analysis, we do not intend to linearise about $\tau(z)$, so a difference of any size between this and the ‘true’ steady background state may be included in the perturbation θ .

Using the above form of the temperature field, the equations of heat and momentum

conservation become

$$\frac{\partial \theta}{\partial t} + \mathbf{u} \cdot \nabla \theta = \nabla^2 \theta + \tau'' - w\tau, \quad (42)$$

$$\frac{1}{\text{Pr}} \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) = -\nabla p + \nabla^2 \mathbf{u} + \text{Ra}(\tau + \theta)\hat{\mathbf{z}}, \quad (43)$$

with boundary conditions $\theta(0) = \theta(1) = 0$. Using the identity (18), and denoting by angled brackets a combined spatial/temporal average, we find that

$$\text{Nu} = \langle |\nabla T|^2 \rangle = \int_0^1 [\tau'(z)]^2 dz + 2 \left\langle \tau' \frac{\partial \theta}{\partial z} \right\rangle + \langle |\nabla \theta|^2 \rangle. \quad (44)$$

We can eliminate the term involving $\tau' d\theta/dz$ in (44) by multiplying the heat equation (42) by 2θ taking the time/space average and integrating by parts. The resulting equation is

$$\text{Nu} = \int_0^1 [\tau'(z)]^2 dz - \langle |\nabla \theta|^2 + 2\tau' w\theta \rangle \quad (45)$$

Finally, for convenience later, we add the time/space-average of $2\mathbf{u} \cdot (43)$ to (45) to find that

$$\text{Nu} = \int_0^1 [\tau'(z)]^2 dz - Q\{\theta, \mathbf{u}\}, \quad (46)$$

where

$$Q\{\theta, \mathbf{u}\} = \left\langle |\nabla \theta|^2 + \frac{2}{\text{Ra}} |\nabla \mathbf{u}|^2 + 2(\tau' - 1)w\theta \right\rangle. \quad (47)$$

This general result forms the backbone of the background method. If one can find a particular $\tau(z)$ such that the quadratic form Q is non-negative definite, then the first term on the right-hand side of (46) provides an upper bound on the Nusselt number Nu . An illustration of how one could perform such an analysis is given in detail in the Appendix.

One should not expect that there is a unique $\tau(z)$ that will give a positive definite quadratic form Q . Indeed, one may formulate a variational calculus problem based upon choosing $\tau(z)$ in order to minimise the upper bound subject to the temperature boundary conditions. This very problem has been solved, numerically, by Plasting & Kerswell [27].

It is interesting to note that the background method may be thought of a rigorous implementation of the marginal stability argument proposed in §3.2. For a hypothesised temperature profile (for example, the well-mixed interior with thin boundary layers in Figure 7), we aim to estimate the heat flux obtained as if this profile were an equilibrium solution.

4.2 A brief history of bounds on Nu

The background method is not the only analytical technique to have been applied to the problem of high-Ra convection. The original analysis by Howard [13] used a power balance combined with some statistical hypotheses about the nature of turbulent convection to arrive at the bound $\text{Nu} < C\text{Ra}^{1/2}$ uniformly in Pr . No author since has been able to improve the exponent of Ra for arbitrary (finite) Pr , lending support to the ‘ultimate’ scaling $\text{Nu} \propto \text{Ra}^{1/2}$. The prefactor was improved, however, through the asymptotics of

Busse [3]. This bound was originally unchallenged by the application of the background method Constantin & Doering [5], until Nicodemus *et al.* [20] introduced a so-called ‘balance parameter’ to the background method, allowing the use of largely-isothermal background temperature profiles. It was shown by Kerswell [15] that the best possible bounds obtainable by the background method and by Howard [13] should be one and the same. The most strict currently-known bound was determined by Plasting & Kerswell [27], who obtained the optimal profile for use in the background method, and arrived at the bound

$$\text{Nu} < 0.0264 \text{Ra}^{1/2}. \quad (48)$$

4.3 Convection and heat transport at $\text{Pr} = \infty$

Several authors have extended the the search for a high-Ra scaling to fluids with infinite Prandtl number - that is, fluids with very small thermal diffusivity, relative to their (kinematic) viscosity. A very important example of such a fluid is the Earth’s mantle. In the infinite-Pr regime the equations of heat, momentum and mass conservation are

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = \nabla^2 T, \quad (49)$$

$$\nabla p = \nabla^2 \mathbf{u} + \text{Ra} T \hat{\mathbf{z}}, \quad (50)$$

$$\nabla \cdot \mathbf{u} = 0. \quad (51)$$

Note that the velocity and pressure fields are slaved to the temperature field T in this case. One can perform both a linear stability analysis and an energy analysis in exactly the same manner as for the finite Pr fluid case illustrated in §§2.2-2.3, arriving at precisely the same result. The flow is linearly unstable for $\text{Ra} > \text{Ra}_c$ and nonlinearly stable for $\text{Ra} < \text{Ra}_c$, with $\text{Ra}_c = 1708$ for a problem with no-slip boundaries at the top and bottom of the cell, $\text{Ra}_c = 27\pi^4/4$ with stress-free boundaries and so on. Note that, although the marginal stability conditions remain the same, behaviour above Ra_c *does* depend on Pr.

If one were to investigate the behaviour of the Nusselt number Nu at high Ra, however, one finds notable differences between this and the finite Pr case. It is still possible to find analytical upper bounds for Nu using the background method described above; the Nusselt number still takes the form

$$\text{Nu} = \int_0^1 [\tau(z)]^2 dz - Q\{\theta\} \quad (52)$$

but the quadratic form Q becomes

$$Q\{\theta\} = \langle |\nabla \theta|^2 + 2\tau' w \theta \rangle, \quad \text{where} \quad \nabla^2 w = -\text{Ra} \nabla_H^2 \theta. \quad (53)$$

Analyses have yielded a number of bounds on Nu , each specifying different power laws. The first application of the background method to this problem Doering & Constantin [9] yielded a bound of the form $\text{Nu} \leq C\text{Ra}^{2/5}$. It was later proved, both by Otero [23] and by Plasting [26], that $\text{Ra}^{2/5}$ is the best possible scaling that may be obtained with a monotone background $\tau(z)$, using the background method alone. Yan [33] improved the bound to $\text{Ra}^{4/11}$ by instead using a maximum principle applied to a monotone profile. The addition of a singular integral analysis to this maximum principle can yield a scaling of the form

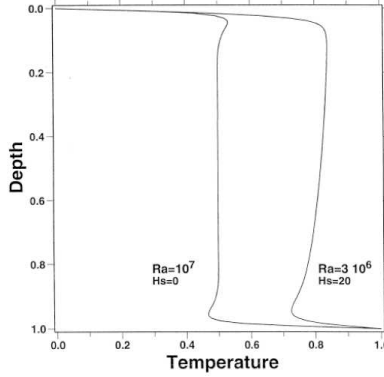


Figure 8: The non-monotonic average temperature profile found by Sotin & Labrosse [30].

$Nu \leq cRa^{1/3}(\log Ra)^{2/3}$ [6]. More recently, Doering, Otto & Reznikoff [10] further refined the rigorous bound to $(Ra \log Ra)^{1/3}$ by applying the background method with a non-monotone background function. A numerical computation of the optimal background profile (which turns out to be non-monotone), suggests a simple $Ra^{1/3}$ scaling for the Nu bound [14].

The need for a non-monotonic background function for the temperature was anticipated by the observations of Sotin & Labrosse [30], who carried out a detailed numerical simulation of infinite Prandtl number convection. In their simulations², it was apparent that cold fluid, after having descended close to the base of the convection cell, ‘pools’ above the thermal boundary layer due to heating of the lower boundary. This results in a small depression of mean temperature between the edge of the lower boundary layer and the well-mixed central region. The same phenomenon occurs with hot fluid just below the upper boundary layer. This ‘lid’ on the boundary layer stabilizes it, allowing it to grow nearly to $Ra^{-1/3}$ as conjectured by the marginal stability argument given in §3.2. The resulting time-averaged temperature profiles resemble those shown in Figure 4.3.

4.4 Bounds on convection in porous media

One can also investigate convection in a porous medium, where the fluid flow is governed by Darcy’s law

$$\nabla p = -\mathbf{u} + Ra_D T \hat{\mathbf{z}}, \quad \nabla \cdot \mathbf{u} = 0, \quad (54)$$

with the Darcy-Rayleigh number

$$Ra_D = \frac{g\alpha\Delta T h K}{\nu\kappa}. \quad (55)$$

K here represents the *permeability* of the medium (the square of the lengthscale of the pores, essentially a measure of the mobility of fluid within the porous matrix). In this problem, at least, there is much less controversy over the scaling of the Nusselt number for large Ra_D . Analogous arguments to those used to get the $Ra^{1/2}$ and $Ra^{1/3}$ scalings for

²viewable online at the time of writing at <http://www.ipgp.jussieu.fr/~labrosse/movies.html>

open fluids both suggest that $\text{Nu} \propto \text{Ra}_D^1$ as $\text{Ra}_D \rightarrow \infty$. This statement is further supported by an application of the background method, such as has been performed by Doering & Constantin [8].

It should be noted, however, that the scaling $\text{Nu} \propto \text{Ra}_D$ is rarely observed in experiments. Indeed, it is difficult to extract a single scaling law from the experimental data, as compiled by Lister [17]. Part of the problem lies in attempting to define a global Darcy-Rayleigh number on what is inevitably a heterogeneous medium. Furthermore, Darcy's law itself can be called into question, depending upon the physical parameter used to increase Ra_D . See also Otero *et al.* [24] for a detailed discussion of the analysis of convection in a porous medium.

5 Open problems and challenges

We conclude with a small subset of the many open problems and challenging aspects not yet fully explored in this interesting field.

Given the numerous inconsistent experimental results for the scaling of Nu at high Ra , one would very much like to know the ultimate state of Rayleigh-Bénard convection. Does a single ultimate state exist, or are different states reached, depending on the specify geometry or region of parameter space, as suggested by Grossmann & Lohse [12]?

Bounds on Nu derived via the background method do not yet include details of high- Ra dependence of Nu on the Prandtl number Pr . Is it possible to obtain sharp, uniform bounds on the Nusselt number when considering both dimensionless parameters?

Several generalisations of simple Rayleigh-Bénard convection are of interest to geophysicists, in the context of thermal transport in the mantle, or in the oceans and atmosphere. We could ask about the comparison between fixed temperature and fixed heat flux boundary conditions (see [25]), or examine the effect of adding rotation in the form of a Coriolis force (see [7], for example).

Finally, problems involving free boundaries present an interesting variation on convection problems, largely neglected in the current literature (except of course for the original closed-form linear solution of Rayleigh [28]). Not only would stress-free boundary conditions³ alter the analysis leading to upper bounds, but more Marangoni effects, due to the variation of surface tension with temperature, could present some interesting results. Indeed, this would bring us back to our starting point, with the observations of Bénard [2].

A Derivation of integral identities for Nu

Integrating equation (15) over the domain V , we obtain

$$\frac{d}{dt} \int_V T dV + \int_{\partial V} T \mathbf{u} \cdot \mathbf{dS} = \int_{\partial V} \nabla T \cdot \mathbf{dS}, \quad (56)$$

³[23] has conjectured, based on a numerical implementation of the background method, that $\text{Nu} \lesssim \text{Ra}^{5/12}$ for two-dimensional convection with stress-free boundaries

where we have used the fact that $\nabla \cdot \mathbf{u} = 0$. After applying our boundary conditions to the two surface integrals, we are left with

$$\frac{d}{dt} \int_V T dV = \int_{z=1} \frac{\partial T}{\partial z} dS - \int_{z=0} \frac{\partial T}{\partial z} dS. \quad (57)$$

Similarly, after multiplying equation (15) by T and integrating over V , we obtain

$$\frac{d}{dt} \int_V \frac{1}{2} T^2 dV + \int_{\partial V} \frac{1}{2} T^2 \mathbf{u} \cdot \mathbf{dS} = \int_{\partial V} T \nabla T \cdot \mathbf{dS} - \int_V |\nabla T|^2 dV \quad (58)$$

$$\Rightarrow \frac{d}{dt} \int_V \frac{1}{2} T^2 dV = \int_{z=0} -\frac{\partial T}{\partial z} dS - \int_V |\nabla T|^2 dV. \quad (59)$$

Finally, multiplying equation (15) by z and integrating over V , we obtain

$$\frac{d}{dt} \int_V zT dV + \int_{\partial V} zT \mathbf{u} \cdot \mathbf{dS} - \int_V wT dV = \int_{\partial V} z \nabla T \cdot \mathbf{dS} - \int_V \frac{\partial T}{\partial z} dV \quad (60)$$

$$\Rightarrow \frac{d}{dt} \int_V zT dV - \int_V wT dV = \int_{z=1} \frac{\partial T}{\partial z} dS + \int_{z=0} dS. \quad (61)$$

We now put (57), (59) and (61) together to find

$$\frac{d}{dt} \int_V (T - \frac{1}{2} T^2 - zT) dV + \int_V wT dV = \int_V |\nabla T|^2 dV - \int_{z=0} dS. \quad (62)$$

We define the time average operator $\bar{\cdot}$ by

$$\bar{f} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t=0}^{t=T} f(t) dt \quad (63)$$

and note that, if $f(t)$ is bounded, then $\frac{df}{dt}$ vanishes⁴. Assuming that T remains bounded throughout the domain, the time average of (62) then yields

$$\overline{\int_V wT dV} = \overline{\int_V |\nabla T|^2 dV} - \overline{\int_{z=0} dS} \quad (64)$$

$$\Rightarrow \langle wT \rangle = \langle |\nabla T|^2 \rangle - 1. \quad (65)$$

Furthermore, if we integrate equation (23) over the domain, we find

$$\frac{1}{\text{Pr}} \frac{d}{dt} \int_V \frac{1}{2} u^2 dV = - \int_V |\nabla \mathbf{u}|^2 dV + \text{Ra} \int_V Tw dV \quad (66)$$

and so

$$0 = - \langle |\nabla \mathbf{u}|^2 \rangle + \text{Ra} \langle Tw \rangle. \quad (67)$$

⁴If, say, $|f(t)| \leq C < \infty \forall t$ then $\frac{1}{T} \left| \int_0^T \frac{df}{dt}(t) dt \right| = \frac{1}{T} |f(T) - f(0)| \leq \frac{2C}{T} \rightarrow 0$ as $T \rightarrow \infty$.

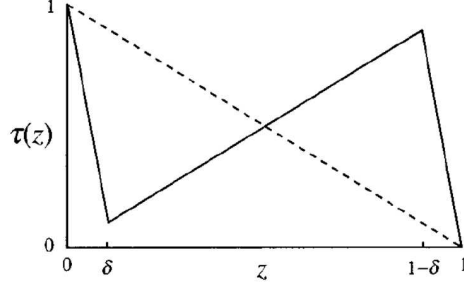


Figure 9: The piecewise linear background function (71) used in the derivation of an upper bound for Nu.

B A sample application of the background method

In this appendix, we give an example of the application of the background method to a simple case, in order to obtain a relatively crude upper bound for Nu. As in the main body of the notes (see §4.1), we write the temperature field in the form

$$T(\mathbf{x}, t) = \tau(z) + \theta(\mathbf{x}, t) \quad (68)$$

and consider the magnitude of the Nusselt number in the form

$$\text{Nu} = \int_0^1 [\tau'(z)]^2 dz - Q\{\theta, \mathbf{u}\}, \quad (69)$$

where the quadratic form Q is given by

$$Q\{\theta, \mathbf{u}\} = \left\langle |\nabla \theta|^2 + \frac{2}{\text{Ra}} |\nabla \mathbf{u}|^2 + 2(\tau' - 1)w\theta \right\rangle. \quad (70)$$

We would now like to find a lower bound for Q , giving us an upper bound for Nu. Each of these bounds will depend upon the form of the background function $\tau(z)$. Whilst the first two terms in Q are positive definite, we can say very little about the last. Using our freedom to choose the background temperature field, we use a piecewise linear background $\tau(z)$ illustrated in Figure 9, and defined mathematically by

$$\tau(z) = \begin{cases} 1 - \left(\frac{1-\delta}{\delta}\right) z & \text{for } 0 < z < \delta, \\ z & \text{for } \delta < z < 1 - \delta, \\ \left(\frac{1-\delta}{\delta}\right) (1 - z) & \text{for } 1 - \delta < z < 1, \end{cases} \quad (71)$$

which eliminates the final term in Q everywhere except within a region of width δ at each boundary. For this $\tau(z)$, we note that

$$\int_0^1 [\tau'(z)]^2 dz = \frac{2}{\delta} - 3, \quad (72)$$

before concentrating our attention on bounding the quadratic form $Q\{\theta, \mathbf{u}\}$.

The absolute value of the last term in (70) is given for the background temperature field (71) by

$$\left| \int_V (\tau'(z) - 1) w \theta dV \right| = \frac{1}{\delta} \left| \int \int \left(\int_0^\delta w \theta dz \right) dy dx + \int \int \left(\int_{1-\delta}^1 w \theta dz \right) dy dx \right|. \quad (73)$$

Using a combination of the fundamental theorem of calculus and the Cauchy-Schwarz inequality, we can bound the temperature perturbation as follows

$$|\theta(x, y, z)| = \left| \int_0^z \frac{\partial \theta}{\partial z}(x, y, \zeta) d\zeta \right| \quad (74)$$

$$\leq \left(\int_0^z d\zeta \right)^{1/2} \left(\int_0^z \left(\frac{\partial \theta}{\partial z} \right)^2 d\zeta \right)^{1/2} \quad (75)$$

$$\leq z^{1/2} \left(\int_0^{1/2} \left(\frac{\partial \theta}{\partial z} \right)^2 d\zeta \right)^{1/2} \quad (\text{if } z \leq 1/2). \quad (76)$$

After having found a similar bound for the vertical velocity w , and using a similar argument for $1/2 \leq z \leq 1$, we can now bound the integrals in (73) in the following way

$$\frac{1}{\delta} \left| \int \int \left(\int_0^\delta w \theta dz \right) dy dx \right| \leq \frac{\delta}{2} \int \int \left\{ \left(\int_0^{1/2} \left(\frac{\partial \theta}{\partial z} \right)^2 d\zeta \right)^{1/2} \left(\int_0^{1/2} \left(\frac{\partial w}{\partial z} \right)^2 d\zeta \right)^{1/2} \right\} dx dy \quad (77)$$

$$\leq \frac{\delta^2 \text{Ra}}{16} \int_{z \leq 1/2} \left(\frac{\partial \theta}{\partial z} \right)^2 dV + \frac{1}{\text{Ra}} \int_{z \leq 1/2} \left(\frac{\partial w}{\partial z} \right)^2 dV. \quad (78)$$

In obtaining the above bound, we have used the inequality $2ab \leq a^2 + b^2$, with a and b carefully chosen so that the velocity gradient term will have the same prefactor ($2/\text{Ra}$) as in the expression for $Q\{\theta, \mathbf{u}\}$ (70). After combining the contributions from the boundary layers near 0 and 1, we can write

$$|\langle 2(\tau'(z) - 1)w\theta \rangle| \leq \frac{\delta^2 \text{Ra}}{8} \int_V \left(\frac{\partial \theta}{\partial z} \right)^2 dV + \frac{2}{\text{Ra}} \int_V \left(\frac{\partial w}{\partial z} \right)^2 dV, \quad (79)$$

therefore we can bound the quadratic form Q from below as follows

$$Q\{\theta, \mathbf{u}\} = \left\langle |\nabla \theta|^2 + \frac{2}{\text{Ra}} |\nabla \mathbf{u}|^2 + 2(\tau' - 1)w\theta \right\rangle \quad (80)$$

$$\geq \left\langle \left(\frac{\partial \theta}{\partial z} \right)^2 + \frac{2}{\text{Ra}} \left(\frac{\partial w}{\partial z} \right)^2 - \frac{\delta \text{Ra}}{8} \left(\frac{\partial \theta}{\partial z} \right)^2 - \frac{2}{\text{Ra}} \left(\frac{\partial w}{\partial z} \right)^2 \right\rangle \quad (81)$$

$$= \left(1 - \frac{\delta^2 \text{Ra}}{8} \right) \left\langle \left(\frac{\partial \theta}{\partial z} \right)^2 \right\rangle. \quad (82)$$

If we use our freedom to choose the background function so that $\delta = 2\sqrt{2/\text{Ra}}$, we can guarantee that $Q \geq 0$ for all θ and \mathbf{u} . Substituting this information into equations (71) and (72), we find that

$$\text{Nu} \leq \sqrt{\frac{\text{Ra}}{2}} - 3. \quad (83)$$

For consistency, we require that $\delta \leq 1/2$, as the background function must be single-valued. The above analysis therefore only holds for $\text{Ra} \geq 32$.

References

- [1] G. AMATI, K. KOAL, F. MASSAIOLI, K. R. SREENIVASAN, AND R. VERZICCO, *Turbulent thermal convection at high rayleigh numbers for a boussinesq fluid of constant prandtl number*, Phys. Fluids, 17 (2005), p. 121701.
- [2] H. BÉNARD, *Let tourbillons cellulaires dans une nappe liquide*, Rev. Gén. Sci. Pures Appl., 11 (1900), pp. 1261–1271.
- [3] F. H. BUSSE, *On howard’s upper bound for heat transport by thermal convection*, J. Fluid Mech., 37 (1969), pp. 457–477.
- [4] X. CHAVANNE, F. CHILLÀ, B. CASTAING, B. HÉBRAL, B. CHABAUD, AND J. CHAUSSY, *Observation of the ultimate regime in rayleigh-bénard convection*, Phys. Rev. Lett., 79 (1997), pp. 3648–3651.
- [5] P. CONSTANTIN AND C. R. DOERING, *Variational bounds on energy dissipation in incompressible flows. iii. convection*, Phys. Rev. E, 53 (1996), pp. 5957–5981.
- [6] ———, *Infinite prandtl number convection*, J. Stat. Phys., 94 (1999), pp. 159–172.
- [7] P. CONSTANTIN, C. HALLSTROM, AND V. POUTKARADZE, *Logarithmic bounds on infinite prandtl number rotating convection*, J. Math. Phys., 42 (2001), p. 773.
- [8] C. R. DOERING AND P. CONSTANTIN, *Bounds for heat transport in a porous layer*, J. Fluid Mech., 376 (1998), pp. 263–296.
- [9] ———, *On upper bounds for the infinite prandtl number convection with or without rotation*, J. Math. Phys., 42 (2001), pp. 784–795.
- [10] C. R. DOERING, F. OTTO, AND M. G. REZNIKOFF, *Bounds on vertical heat transport for infinite-prandtl-number rayleigh-bénard convection*, J. Fluid Mech., 560 (2006), pp. 229–241.
- [11] J. A. GLAZIER, T. SEGAWA, A. NAERT, AND M. SANO, *Evidence against ‘ultrahard’ thermal turbulence at very high rayleigh numbers*, Nature, 398 (1999), pp. 307–310.
- [12] S. GROSSMANN AND D. LOHSE, *Scaling in thermal convection: a unifying theory*, J. Fluid Mech., 407 (2000), pp. 27–58.

- [13] L. N. HOWARD, *Heat transport by turbulent convection*, J. Fluid Mech., 17 (1963), p. 405.
- [14] G. R. IERLEY, R. R. KERSWELL, AND S. C. PLASTING, *Infinite-prandtl-number convection. part 2. a singular limit of upper bound theory*, J. Fluid Mech., 560 (2006), pp. 159–227.
- [15] R. R. KERSWELL, *Variational bounds on shear-driven turbulence and turbulent boussinesq convection*, Physica D, 100 (1997), pp. 355–376.
- [16] R. H. KRAICHNAN, *Turbulent thermal convection at arbitrary prandtl number*, Phys. Fluids, 5 (1962), pp. 1374–1389.
- [17] C. R. B. LISTER, *An explanation for the multivalued heat transport found experimentally for convection in a porous medium*, J. Fluid Mech., 214 (1990), pp. 287–320.
- [18] W. V. R. MALKUS, *Discrete transitions in turbulent convection*, Proc. Roy. Soc. Lond. A, 225 (1954), p. 185.
- [19] ———, *The heat transport and spectrum of thermal turbulence*, Proc. Roy. Soc. Lond. A, 225 (1954), p. 196.
- [20] R. NICODEMUS, S. GROSSMANN, AND M. HOLTHAUS, *Improved variational principle for bounds on energy dissipation in turbulent shear flow*, Physica D, 101 (1997), pp. 178–190.
- [21] J. J. NIEMELA, L. SKRBEK, K. R. SREENIVASAN, AND R. J. DONNELLY, *Turbulent convection at very high rayleigh numbers*, Nature, 404 (2000), pp. 837–840.
- [22] A. NIKOLAENKO, E. BROWN, D. FUNFSCHILLING, AND G. AHLERS, *Heat transport by turbulent rayleigh-bénard convection in cylindrical cells with aspect ratio one and less*, J. Fluid Mech., 523 (2005), pp. 251–260.
- [23] J. OTERO, *Bounds for the heat transport in turbulent convection*, PhD thesis, University of Michigan, 2002.
- [24] J. OTERO, L. A. DONTCHEVA, H. JOHNSTON, R. A. WORTHING, A. KURGANOV, G. PETROVA, AND C. R. DOERING, *High-rayleigh-number convection in a fluid-saturated porous layer*, J. Fluid Mech., 500 (2004), pp. 263–281.
- [25] J. OTERO, R. W. WITTENBERG, R. A. WORTHING, AND C. R. DOERING, *Bounds on rayleigh-bénard convection with an imposed heat flux*, J. Fluid Mech., 473 (2002), pp. 191–199.
- [26] S. C. PLASTING, *Turbulence has its limits: a priori estimates of transport properties in turbulent fluid flows*, PhD thesis, University of Bristol, 2004.
- [27] S. C. PLASTING AND R. R. KERSWELL, *Improved upper bound on the energy dissipation rate in plane couette flow: the full solution to busse’s problem and the constantin-doering-hopf problem with one-dimensional background field*, J. Fluid Mech., 477 (2003), pp. 363–379.

- [28] L. RAYLEIGH, *On convection currents in a horizontal layer of fluid, when the higher temperature is on the under side*, Phil. Mag., 32 (1916), pp. 529–538.
- [29] P.-E. ROCHE, B. CASTAING, B. CHABAUD, AND B. HÉBRAL, *Observation of the $1/2$ power law in rayleigh-bénard convection*, Phys. Rev. E, 63 (2001), p. 045303.
- [30] C. SOTIN AND S. LABROSSE, *Three-dimensional thermal convection in an iso-viscous, infinite prandtl number fluid heated from within and from below: applications to the transfer of heat through planetary mantles*, Phys. Earth Plan. Interiors, 112 (1999), pp. 171–190.
- [31] E. A. SPIEGEL, *Convection in stars i. basic boussinesq convection*, Ann. Rev. Astronomy Astrophys., 9 (1971), pp. 323–352.
- [32] G. STRINGANO, G. PASCAZIO, AND R. VERZICCO, *Turbulent thermal convection over grooved plates*, J. Fluid Mech., 557 (2006), pp. 307–336.
- [33] X. YAN, *On limits to convective heat transport at infinite prandtl number with or without rotation*, J. Math. Phys., 45 (2004), p. 2718.

Waves and Vortices Driven by Interfacial Fluxes

Lecturer: Kerry A. Emanuel

Notes written by: Malte F. Jansen & Yutian Wu

November 17, 2008

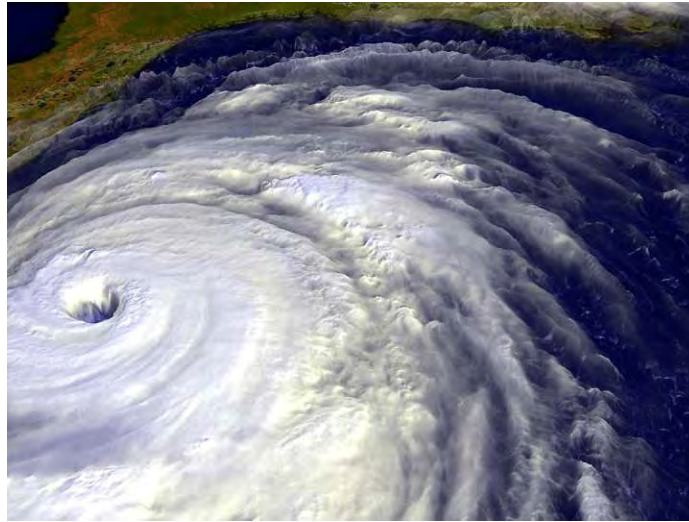


Figure 1: Satellite image of Hurricane Floyd approaching the east coast of Florida in 1999. The image has been digitally enhanced to lend a three-dimensional perspective. Credit: NASA/Goddard Space Flight Center.

1 Atmospheric Boundary Layers

The flux of sensible and latent heat at the surface boundary layer can drive major atmospheric disturbances as for example hurricanes. On the other hand the surface winds associated with these disturbances, influence the surface fluxes which leads to an essential feedback. Our understanding of boundary layer fluxes is therefore crucial for the understanding of many meteorological phenomena. We will here derive some simple scaling laws for the simplified cases of a boundary below a semi-infinite domain. First we will consider the cases of convective and mechanically driven turbulence separately and then briefly discuss how the results might be generalized. In section 1.4 we will show how the interaction between surface winds and surface heat flux can lead to growing wave-like perturbations.

Finally section 1.5 will show that turbulent heat flux in the boundary layer is essential to wipe out the thermodynamically unstable radiative equilibrium state.

1.1 Convective Boundary Layers

Assume a lower boundary under a semi-infinite domain. The turbulent flux of buoyancy through the boundary at $z=0$ is then given by

$$Q = \overline{w'B'} = - \int_0^\infty \dot{B} dz, \quad (1)$$

where \dot{B} denotes the radiative cooling in the domain. The buoyancy flux has the dimension of a length squared over a time cubed:

$$Q \sim L^2 t^{-3} \quad (2)$$

Since no length scale is intrinsic to the system, dimensional analysis gives a velocity scale

$$w' \sim (Qz)^{\frac{1}{3}} \quad (3)$$

i.e. the turbulent velocity increases as $z^{\frac{1}{3}}$. Similarly we find

$$B' \sim Q^{\frac{2}{3}} z^{-\frac{1}{3}} \quad (4)$$

This would imply infinite buoyancy perturbations as $z \rightarrow 0$. To overcome this problem we have to consider a thin surface layer in which diffusion or the roughness of the boundary becomes important. For atmospheric applications the boundary roughness scale is generally much larger than the scale at which molecular diffusion starts to play a role. The above scaling laws then apply only down to the height of the boundary layer roughness scale z_0^T (where T stands for "thermal") and from (3) and (4) we get

$$w'_0 \sim (Qz_0^T)^{\frac{1}{3}} \quad (5)$$

and

$$B'_0 \sim Q^{\frac{2}{3}} z_0^{T-\frac{1}{3}}. \quad (6)$$

Similar to the buoyancy perturbations, the mean buoyancy above the roughness scale has to vary as $z^{-\frac{1}{3}}$. We thus find

$$\bar{B} - \bar{B}_0 \sim Q^{\frac{2}{3}} [(z_0^T)^{-\frac{1}{3}} - z^{-\frac{1}{3}}]. \quad (7)$$

This result implies that for $z \gg z_0^T$, \bar{B} converges to the constant value $\bar{B} \sim Q^{\frac{2}{3}} z_0^{T-\frac{1}{3}}$ which is determined by the surface flux and the roughness length.

Note that the above arguments change dramatically if an upper lid would be considered. In this case another length scale h given by the height of the domain would exist. The turbulent velocity scale in this case is given by the Deardorff [2] scaling.

$$w' \sim (Qh)^{\frac{1}{3}} \quad (8)$$

If more than one length scale exists in the problem, simple scaling laws can generally no longer be found, since all quantities can be functions of the non-dimensional numbers that are then intrinsic to the problem.

1.2 Shear Driven Boundary Layers

A similar analysis can be performed if turbulence is dominated by mechanical forcing. Again assuming a semi-infinite domain we find that the flux of momentum through the boundary is given by

$$M = \int_0^\infty (\dot{u}) dz \quad (9)$$

where \dot{u} generally represents any velocity source. In a steady state and in a nonrotating system (or if rotation can be neglected on the regarded scales) we find:

$$M = \int_0^\infty \frac{1}{\rho} \frac{\partial p}{\partial x} dz. \quad (10)$$

M has the dimension of a velocity squared. Defining $u^{*2} \equiv M$ we directly find that the turbulent velocity has to scale as

$$w' \sim u^* \quad (11)$$

and is thus constant with height, unlike in the buoyancy driven case. A scaling for the shear of the mean wind is given by

$$\frac{d\bar{u}}{dz} \sim \frac{u^*}{z}. \quad (12)$$

Integrating and assuming that \bar{u} vanishes at z_0 , which is the roughness length for momentum, gives

$$\bar{u} \sim u^* \ln \frac{z}{z_0}. \quad (13)$$

And for the background velocity difference between the heights z_1 and z_2

$$\Delta \bar{u} \sim u^* \ln \frac{z_2}{z_1}. \quad (14)$$

This can be solved to get an equation for the surface momentum flux

$$M = u^{*2} \sim \frac{(\Delta \bar{u})^2}{(\ln \frac{z_2}{z_1})^2}. \quad (15)$$

This equation allows one to determine the surface flux from a given mean velocity profile (given that the necessary proportionality constant is determined).

More generally, scaling analysis show that the surface flux of any tracer θ in a mechanically driven boundary layer is given by

$$F = cu^*(\bar{\theta}_0 - \bar{\theta}_a), \quad (16)$$

where $\bar{\theta}_a$ denotes the background tracer concentration at some fixed height and c is a constant that depends on the choice of this height. For more details on the derivation of this equation see chapter three in [4].

1.3 Boundary Layers with Shear and Convection

In real atmospheric boundary layers, thermal convection and vertical shear of the background flow usually come together. Assuming prescribed surface fluxes of buoyancy and momentum, Q and M , a natural length scale is given by

$$L = \frac{M^{\frac{3}{2}}}{Q_0} = \frac{u^{*3}}{\overline{w'B'}}. \quad (17)$$

This is referred to as the Monin-Obukov length¹. It can be gained by comparing the turbulent vertical velocities in the buoyancy driven boundary layer (3) and in the shear driven boundary layer (11). This makes clear that the Monin Obukov length separates the boundary into regions where the turbulence is dominantly driven by shear ($z \ll L$) and by convection ($z \gg L$). Since this adds a lengthscale to the system, all vertical dependences of quantities can now contain functions of the nondimensional number z/L , which can obviously not be determined from simple scaling laws.

1.4 The Linear WISHE Model

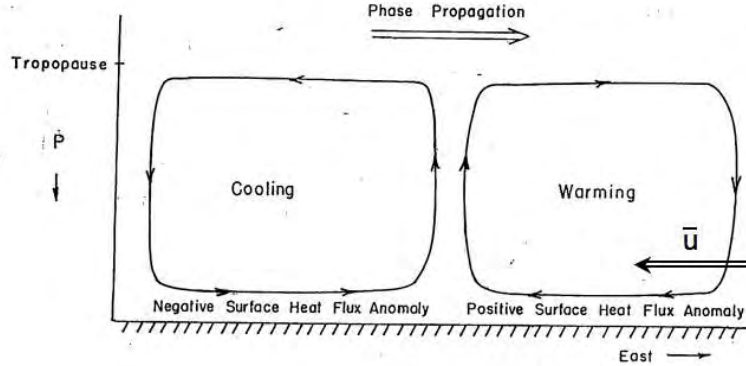


Figure 2: Sketch of the mechanism acting in the WISHE model. Positive superposition of the background easterlies with the easterly disturbance gives rise to an increased surface buoyancy flux, while the partial cancellation of the background flow and the disturbance reduces the buoyancy flux on the left.

The fact that when a momentum source is present, the surface flux of heat depends on the surface wind speed can cause a convective disturbance in the presence of a background flow to grow and propagate, as sketched in figure 2. A simple linear model describing this will be derived in this section. The model is a dry version of the WISHE model derived in [3]. Assuming that the surface heat flux is dominated by mechanically driven turbulence, we can use equation (16) to get the buoyancy flux at the surface

$$Q = c_2 u (B_s - B), \quad (18)$$

¹Strictly speaking the Monin-Obukov length is conventionally defined with the opposite sign, thus being negative if the surface buoyancy flux is positive

where B_s denotes the buoyancy at the surface (which will be assumed to be constant in the following) while B denotes the buoyancy at some fixed finite height above the surface. All variables in (18) denote mean values over a temporal and spatial scale larger than the scale of the background turbulence. In the following we will consider two-dimensional perturbations u' , B' with much larger temporal and spatial scale than the background turbulence and a constant zonal background flow \bar{u} and buoyancy \bar{B} . We further assume that the buoyancy fluctuations are constant with height over the whole depth h of the boundary layer². This is generally a good approximation in a convective layer. Linearizing in the perturbations we then get for the anomalous surface heat flux

$$Q' = h \left(\frac{\partial B'}{\partial t} + \bar{u} \frac{\partial B'}{\partial x} \right) = [c_2 u' (B_s - \bar{B}) - c_2 \bar{u} B']|_{z=0}, \quad (19)$$

where h is the depth of the boundary layer.

The vorticity equation in the x-z plane becomes

$$\left(\frac{\partial}{\partial t} + \bar{u} \frac{\partial}{\partial x} \right) \Delta \psi = \frac{\partial B'}{\partial x} \quad (20)$$

where we used the incompressibility condition which allows us to introduce a streamfunction so that

$$u' = -\frac{\partial \psi}{\partial z} \quad \text{and} \quad w' = \frac{\partial \psi}{\partial x}. \quad (21)$$

Rescaling

$$(x, z) \rightarrow h(x, z) \quad \text{and} \quad t \rightarrow \frac{h}{\bar{u}} t \quad (22)$$

and using (19) and (21), we get

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} \right)^2 \Delta \psi = - \left[\alpha \frac{\partial^2 \psi}{\partial x \partial z} + c_2 \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} \right) \Delta \psi \right]_{z=0} \quad (23)$$

where

$$\alpha = \frac{\bar{Q} h}{\bar{u}^3}.$$

Using an approach of the form $\psi = f(y) \exp(i(kx - ct))$ this can be solved for the phase velocity $c_r = \Re(c)$ and growth rate $\sigma = \Im(c)$. The results for $c_2 = 10^{-3}$ and $\alpha = 1$ as well as $\alpha = 5$ are shown in figure 3. We find exponentially growing disturbances in all cases. The growth rate increases with wavenumber, though the increase stagnates as the wavelength becomes short compared to the depth of the fluid. Further we find that the phase velocity relative to the mean flow is always opposite to the latter. For large α , i.e. small background flows or strong heatflux, and long horizontal wavelength, we even find absolute phase propagation opposite to the mean flow.

²The boundary layer here is in general the convective part of the atmosphere up to the first inversion. In the tropics this would usually be the tropopause

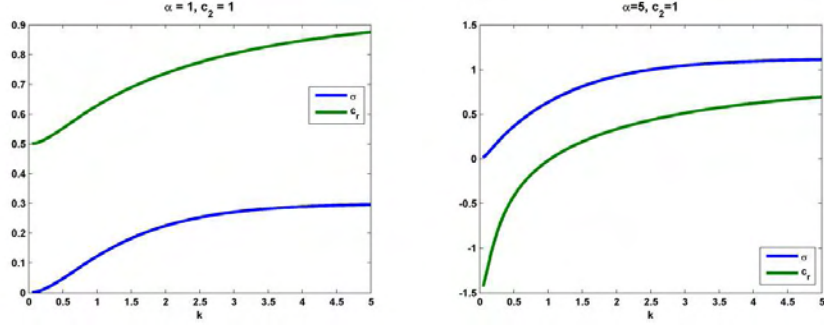


Figure 3: Left: Nondimensional growth rate (blue) and phase propagation (green) for the WISHE model with $\alpha = 1$ and $c_2 = 10^{-3}$. Right: As left but for $\alpha = 5$. Note that the mean flow has velocity 1, i.e. the phase propagation relative to the mean flow is always negative.

1.5 Radiative Equilibrium and Thermodynamic Disequilibrium

Convective adjustment is an important process in the atmosphere, because radiative equilibrium temperature profiles are generally in thermodynamic disequilibrium with the surface. This shall be illustrated here with a simple layered radiative equilibrium model for the atmosphere.

The model is sketched in figure 4. The incoming shortwave radiation is described by an equivalent effective emission temperature T_e . The atmosphere consists of one layer which is completely opaque ($\epsilon = 1$) for longwave radiation but transparent to short-wave radiation. An additional thin layer ($\epsilon_A \ll 1$) represents the atmospheric boundary just above the surface. The radiative equilibrium for the whole system directly yields that

$$T_1 = T_e. \quad (24)$$

From the radiative equilibrium for the surface and the boundary layer, and neglecting the radiative effects of this thin surface layer, layer we find

$$T_s^4 = T_e^4 + T_1^4 \quad (25)$$

and

$$2T_A^4 = T_1^4 + T_s^4. \quad (26)$$

This yields

$$T_A^4 = \left(\frac{3}{4}\right) T_s^4 \quad (27)$$

and therefore $T_A < T_s$. We thus find that in radiative equilibrium the temperature of the atmosphere right above the surface would be lower than the surface temperature itself which would be instable and trigger convection. Qualitatively similar result are obtained with more complex continuous radiative equilibrium models. Thus radiative equilibrium states will generally have convecting boundary layers.

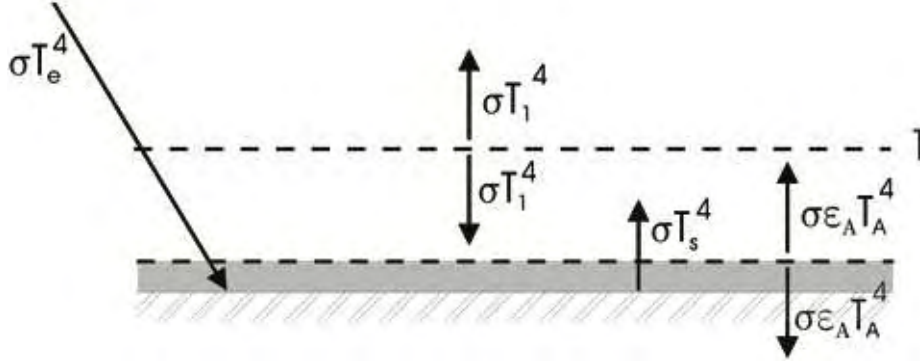


Figure 4: Sketch of a radiative equilibrium model. See text for details.

2 Compressible Convection

For a homogeneous fluid, the specific volume α ($=1/\rho$) can be expressed as a function of two variables, which are chosen to be pressure p and entropy s_d here, i.e. $\alpha = \alpha(p, s_d)$. The entropy (in dry case) s_d is defined by:

$$s_d = c_p \ln\left(\frac{T}{T_o}\right) - R_d \ln\left(\frac{p}{p_o}\right) \quad (28)$$

where T_o and p_o are reference temperature and pressure. We can derive buoyancy B based on entropy s_d :

$$\begin{aligned} B &= -g \frac{(\delta\rho)_p}{\rho} \\ &= g \frac{(\delta\alpha)_p}{\alpha} \\ &= -\frac{dp}{dz} (\delta\alpha)_p \quad (\text{by hydrostatic balance}) \end{aligned} \quad (29)$$

(30)

then, since

$$(\delta\alpha)_p = \left(\frac{\partial\alpha}{\partial s_d}\right)_p \delta s_d = \left(\frac{\partial T}{\partial p}\right)_{s_d} \delta s_d, \quad (31)$$

where the last equation is because of the Maxwell equation:

$$\left(\frac{\partial\alpha}{\partial s_d}\right)_p = \left(\frac{\partial T}{\partial p}\right)_{s_d} \quad (32)$$

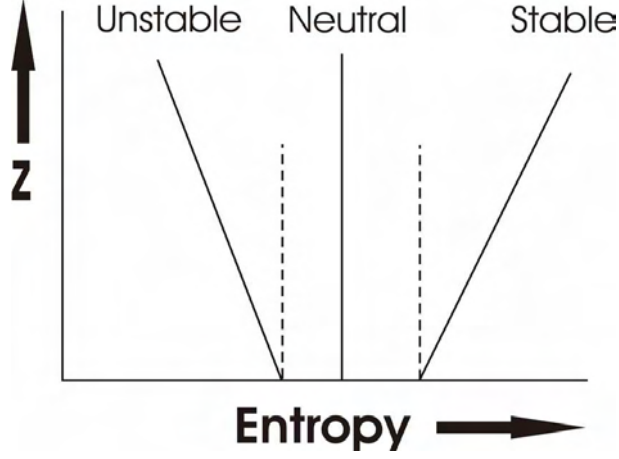


Figure 5: Entropy s_d as a function of height z .

$$\begin{aligned}
 B &= -\frac{dp}{dz} \left(\frac{\partial T}{\partial p} \right)_{s_d} \delta s_d \quad (\text{plug equation 31 into 30}) \\
 &= -\left(\frac{\partial T}{\partial z} \right)_{s_d} \delta s_d
 \end{aligned} \tag{33}$$

$$\equiv \Gamma \delta s_d \tag{34}$$

where $\Gamma \equiv -\left(\frac{\partial T}{\partial z} \right)_{s_d}$ = dry adiabatic lapse rate (rate of change of T for a parcel of air moved vertically and adiabatically) = $\frac{g}{c_p} = 9.8K/km$ for the Earth's atmosphere.

The potential temperature θ , a measure of the entropy of a gas, is defined by

$$\theta = T \left(\frac{p_o}{p} \right)^{R/c_p} \tag{35}$$

where p is the pressure, p_o some reference pressure, R the ideal gas constant, and c_p the heat capacity of the gas at constant pressure. The relation between θ and s_d is given by:

$$s_d = c_p \ln(\theta) + \text{constant}. \tag{36}$$

Both variables are conserved for reversible dry adiabatic processes. Figure 5 shows the stability of entropy as a function of height. It is unstable if entropy decreases with height while it is stable if entropy increases with height. A model aircraft measurement in a desert region near Albuquerque, New Mexico, on August 1993 is shown in Figure 6. In this dry case, the virtual potential temperature doesn't change much with height and the mean entropy is conserved [7].

However, above a thin boundary layer, most atmospheric convection involves phase change of water. Moist convection has significant heating owing to phase changes of water, and helps global redistribution of water vapor. Water vapor, being one of the most important greenhouse gases, is the primary tropospheric infrared absorber while condensed phases are a significant contributor to stratiform cloudiness which absorb infrared radiation and scatter short wave radiation as well. Figure 7 shows the phase equilibria of H_2O , and changes of phase cause significant release or absorption of heat.

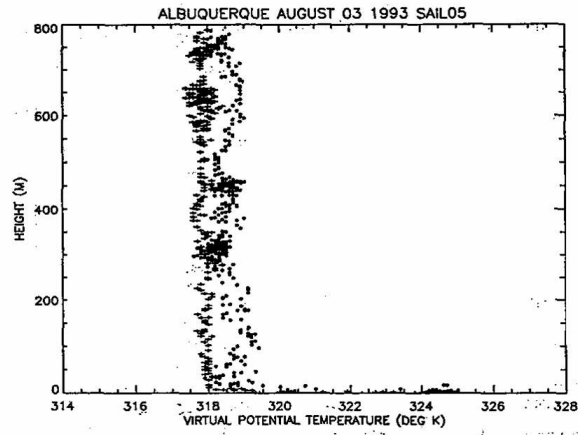


Figure 6: Observations of virtual potential temperature made with an RPV (remotely piloted vehicle) flight in a desert region near Albuquerque, New Mexico, on August 1993. The observations during ascent (inside the plume) are represented by dots; observations on descent (outside of the plume) are represented by crosses.

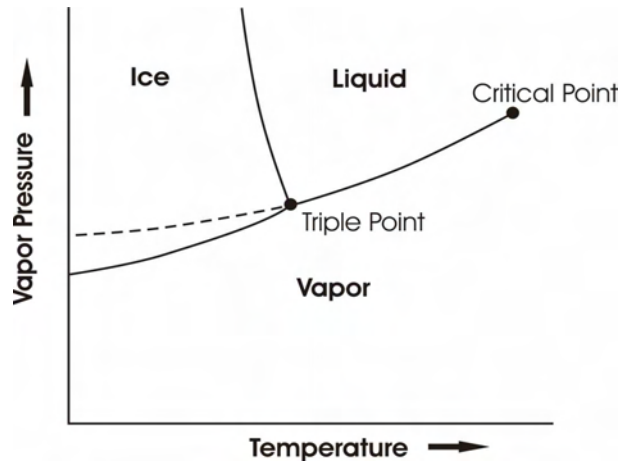


Figure 7: Phase equilibria of H_2O .

The equation of state for water vapor is given by:

$$e\alpha_v = R_v T \quad (37)$$

where e is the vapor pressure, α_v the specific volume for water vapor, and $R_v = R \frac{\bar{m}}{m_v}$ the gas constant for water vapor.

$$e = \frac{R_v T}{\alpha_v} = \frac{R \frac{\bar{m}}{m_v} T}{\alpha_v} = qp \frac{\bar{m}}{m_v} \quad (38)$$

gives the expression for vapor pressure, where q is the mass concentration of water vapor, \bar{m} the mean molecular weight of air, m_v the molecular weight for water vapor. Saturation vapor pressure e^* is also a function of T , i.e. $e^* = e^*(T)$. If we decrease p (or q), both e and e^* will decrease, but since e^* decreases more rapidly, i.e. $e > e^*$, water vapor will be supersaturated.

Supersaturated vapor always condenses onto ambient aerosols called cloud condensation nuclei in a process called heterogeneous nucleation. Cloud condensation nuclei include sea salt from bursting bubbles, windblown dust, combustion products (i.e., organic carbon and soot), photochemically-produced sulfate (i.e., smog), volcanic aerosols, and meteoric debris. The drop size distribution is quite sensitive to the size distribution of cloud condensation nuclei. While stochastic coalescence is a way to form precipitation, the Bergeron-Findeisen process is another important alternative way of initiating precipitation in mid-latitude clouds. Precipitation formation is a strongly nonlinear function of the cloud water concentration. The time scale of precipitation formation is about 10 to 30 minutes.

We used entropy s_d to discuss the stability for dry air, and specific volume α can be expressed as a function of s_d and p , i.e. $\alpha = \alpha(s_d, p)$. In the moist convection case, we can define an approximately conserved thermodynamic variable, the specific entropy s , a function of temperature, pressure, and water concentration,

$$s = c_p \ln\left(\frac{T}{T_o}\right) - R_d \ln\left(\frac{p}{p_o}\right) + L_v \frac{q}{T} - q R_v \ln(H) \quad (39)$$

where $H \equiv \frac{e}{e^*}$ = relative humidity. Specific volume α now depends on three variables rather than two, i.e. $\alpha = \alpha(s, p, q_t)$, where q_t is the total concentration of H_2O of all phases, thus we cannot compare the densities of two samples at the same pressure using just a single entropy variable, a fact of profound consequences for the character of moist convection. But we can make progress by first defining a saturation entropy s^* , the specific entropy air would have if it were saturated with water vapor at the same temperature and pressure. It is given by:

$$s^* = c_p \ln\left(\frac{T}{T_o}\right) - R_d \ln\left(\frac{p}{p_o}\right) + \frac{L_v q^*}{T} = s(T, p, q^*) \quad (40)$$

and

$$\alpha = \alpha(s^*, p, q_t) \quad (41)$$

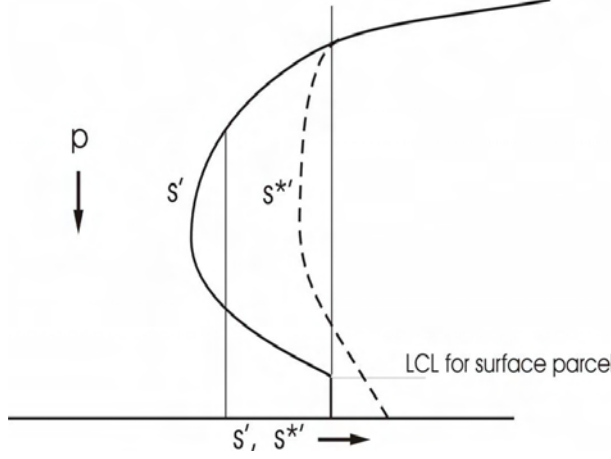


Figure 8: Structure of s' and $s^{*'}.$

Then we can add an arbitrary function of q_t to s^* such that the specific volume can be expressed approximately by two variables,

$$\alpha \cong \alpha(s^{*'}, p) \quad (42)$$

Figure 8 gives the profile of entropy s' and saturation entropy $s^{*'}.$ of the surrounding atmosphere. The straight upward lines are the entropy profile of a parcel which is lifted upward or downward since entropy is conserved. If a parcel is lifted upward with the same entropy as the surrounding atmosphere at the surface, it will rise dry adiabatic until reaching the lifting condensation level (LCL); further ascent will be saturated adiabatic. However only until reaching the crossing point with $s^{*'}.$ will the parcel be stable; after that the parcel would attain positive buoyancy since its entropy (which will equal its saturation entropy since by then the parcel is saturated) will be larger than that of its environment.

Tropical temperature soundings are, to a first approximation, moist adiabatic. Figure 9 is a buoyancy diagram, showing the difference between the density temperature (a temperature that has been corrected for the presence of water vapor and condensed water, so as to accurately measure the density of the air sample) of a reversibly lifted parcel and that of its environment, as a function of the level from which the parcel is lifted and the level to which it is lifted. Note that the environment is almost neutral to a parcel lifted reversibly from around 950 hPa.

When condensed water precipitates, this irreversible process depletes water from the ascending cloudy currents and causes much of the air to be subsaturated when it subsequently descends. Updrafts due to convective instability can be quite strong over a very small fractional area. After reaching the level of neutral buoyancy, it diverges and spreads out laterally, injecting its properties into the large-scale environment. In equilibrium, the surface enthalpy flux F_k , which is given by $C_k \rho |V| (k_o^* - k_a)$, with k_o^* representing the saturation enthalpy of air in contact with the ocean and k_a the enthalpy at the boundary, should be equal to the vertically integrated radiative cooling, where C_k is the enthalpy exchange coefficient, $|V|$ the magnitude of the surface wind, and the specific enthalpy k given by $k = c_p T + L_v q$.

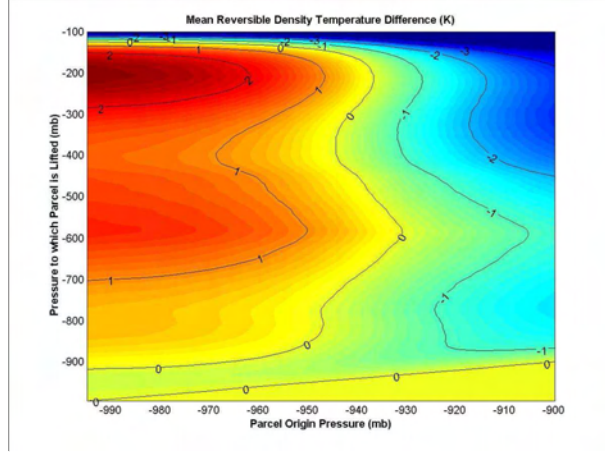


Figure 9: Contour plot of the difference between the environmental density temperature and the density temperature of a parcel lifted reversibly and adiabatically from the pressure level given on the abscissa to the pressure level on the ordinate. This quantity has been averaged over several thousand soundings taken at Kapingamarangi in the tropical western South Pacific.

3 Hurricanes and friends

There are lots of meteorological phenomena related to waves and vortices driven by interfacial fluxes, such as hurricanes, polar lows, dust devils, agukabams, and convectively coupled equatorial waves.

Figure 10 shows the equivalent potential temperature (a measure of moist entropy) of Hurricane Inez in September 28, 1966 as a function of radial distance from the geometrical center of the eye and pressure. Note that the entropy is high at the hurricane center.

By using conserved variables, such as energy, saturation entropy and angular momentum, the potential intensity of the hurricane within the eyewall region can be expressed as:

$$|V|^2 = \frac{C_k}{C_D} \frac{T_s - T_o}{T_o} (k_o^* - k) \quad (43)$$

where C_D is the surface drag coefficient, T_s the sea surface temperature and T_o the temperature at the top of the hurricane [1].

Like the engine in a Carnot Cycle, the energy cycle of hurricanes is one of isothermal expansion, adiabatic expansion, isothermal compression and adiabatic compression. The total rate of heat input to the hurricane is given by:

$$\dot{Q} = 2\pi \int_0^{r_o} \rho [C_k |V| (k_o^* - k) + C_D |V|^3] r dr \quad (44)$$

where r_o is the radius from the storm center, and the first term within the square brackets is the surface enthalpy flux and the second one is the dissipative heating. In steady state, work is used to balance frictional dissipation:

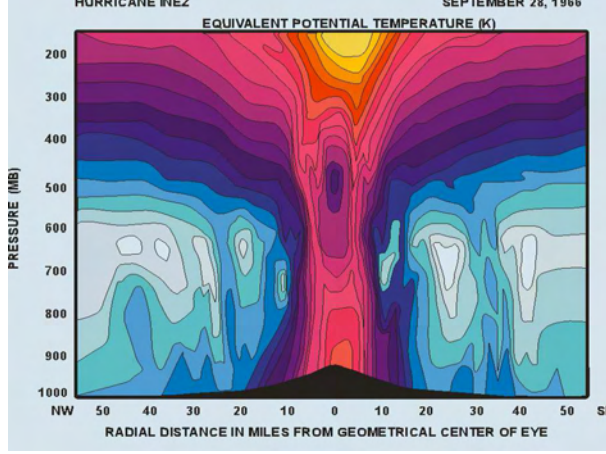


Figure 10: Equivalent potential temperature of Hurricane INEZ in September 28, 1966 as a function of radial distance from the geometrical center of the eye and pressure.

$$W = 2\pi \int_0^{r_o} \rho [C_D |V|^3] r dr \quad (45)$$

$$= \frac{T_s - T_o}{T_s} \dot{Q} \quad (46)$$

$$= 2\pi \frac{T_s - T_o}{T_s} \int_0^{r_o} \rho [C_k |V| (k_o^* - k) + C_D |V|^3] r dr. \quad (47)$$

If we assume that the integrals are dominated by the values of their integrands near the radius of maximum wind speed, then it gives an approximate expression for the maximum wind speed:

$$|V_{max}|^2 \cong \frac{C_k}{C_D} \frac{T_s - T_o}{T_o} (k_o^* - k) \quad [6]. \quad (48)$$

This equation, however, can be derived exactly from considerations of thermal wind balance. Figure 11 shows the maximum wind speed profile as a function of sea surface temperature (SST) and T_o when the relative humidity is 0.75 and C_k/C_D equals 1.2.

Figures 12 and 13 exhibit the relationship between the potential intensity (PI) and intensity of real tropical cyclones [5] [6]. Since the cumulative distribution function (CDF) is linear for tropical storms, within each regime, there is a roughly equal likelihood of a given storm reaching any given intensity up to its potential intensity. In addition, CDFs of the wind speeds in North Atlantic and western North Pacific tropical cyclones were calculated. The peak winds reached during the storms, normalized by their theoretical maximum values, tend to fall into one of two linear CDFs depending on whether the storm does or does not reach hurricane intensity, as shown in Figure 9. This means that a randomly chosen tropical cyclone has the same probability of reaching any given intensity up to marginal hurricane intensity, and another probability of reaching any intensity between marginal hurricane intensity and the maximum theoretical intensity at that time and place.

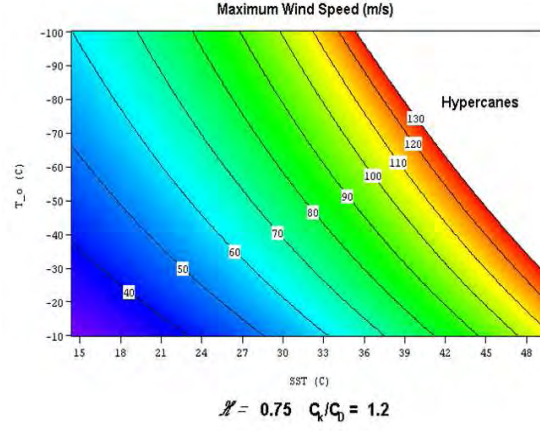


Figure 11: Maximum wind speed profile as a function of sea surface temperature (SST) and T_o when the relative humidity is 0.75 and C_k/C_D equals 1.2.

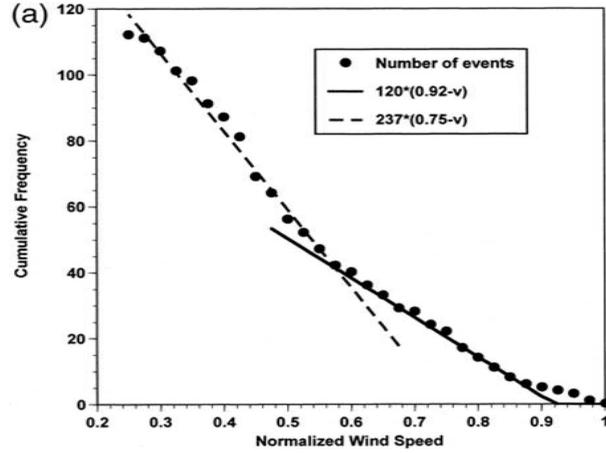


Figure 12: Cumulative distribution function (CDF) of lifetime maximum wind speeds for all tropical cyclones of tropical storm strength 18 m s^{-1} or greater after 1957 whose lifetime maximum intensity was not limited by declining potential intensity. Wind speed is normalized by monthly climatological potential wind speed at the reported position of the tropical cyclones. The ordinate shows the total number of events whose normalized lifetime maximum wind speed exceeds the value on the abscissa.

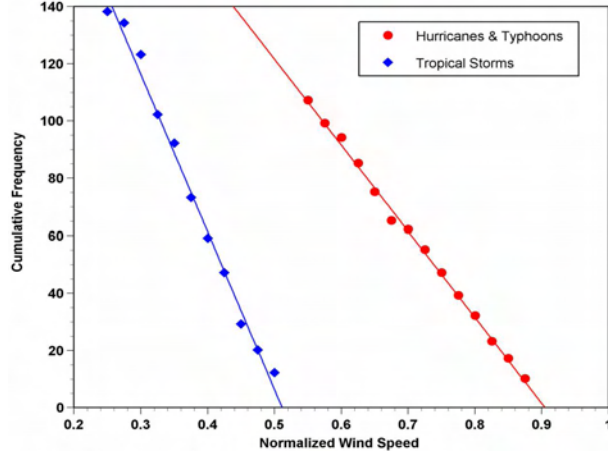


Figure 13: Total number of tropical cyclones with normalized wind speeds exceeding the value on the abscissa, from 1957 to 1999 in the North Atlantic and from 1970 to 1999 in the western North Pacific. Wind speed is normalized by the theoretical maximum wind speed calculated from climatological data.

Polar lows, agukabams, dust devils and convectively coupled equatorial waves are, together with hurricanes, meteorological phenomena related to waves and vortices driven by interfacial fluxes. Polar lows have similar mechanisms as hurricanes, but they are small-scale storms with ice. Their smaller scale is consistent with theoretical estimates of the maximum radius of a convective storm which is proportional to $1/f$. Agukabams draw their energy from enthalpy transferred from hot, moist soils. When agukabams happen, there is usually a big drop in the soil temperature, reflecting the heat transferred from the soil to the air. Dust devils are essentially dry hurricanes, and like agukabams, they draw energy from the sensible heat of the sand. It has been observed that Dust devils happen not only on the Earth but also on Mars.

For convectively coupled equatorial waves, a wavenumber-frequency spectrum analysis of the satellite-observed outgoing longwave radiation (OLR) was performed within the region between $15^{\circ}S - 15^{\circ}N$ by Wheeler and Kiladis (1999) [8]. After removing an estimated background spectrum, the spectral peaks correspond nicely with dispersion relations from equatorially trapped wave modes of shallow water theory. Figure 14 shows the symmetric component of OLR and some peaks corresponding to the dispersion relations of the equatorially trapped wave modes. These so-called convectively coupled equatorial waves are the Kelvin, $n = 1$ equatorial Rossby wave, mixed Rossby-gravity, $n = 0$ eastward inertio-gravity, $n = 1$ westward inertio-gravity (WIG), and $n = 2$ WIG waves. The Madden-Julian oscillation (MJO) and the tropical depression-type (TD-type) disturbances are also present in the spectra, but they are unlike the convectively coupled equatorial waves due to their location away from the equatorial wave dispersion curves in the wavenumber-frequency spectrum.

References

- [1] M. BISTER AND K. A. EMANUEL, *Dissipative heating and hurricane intensity.*, Mete-

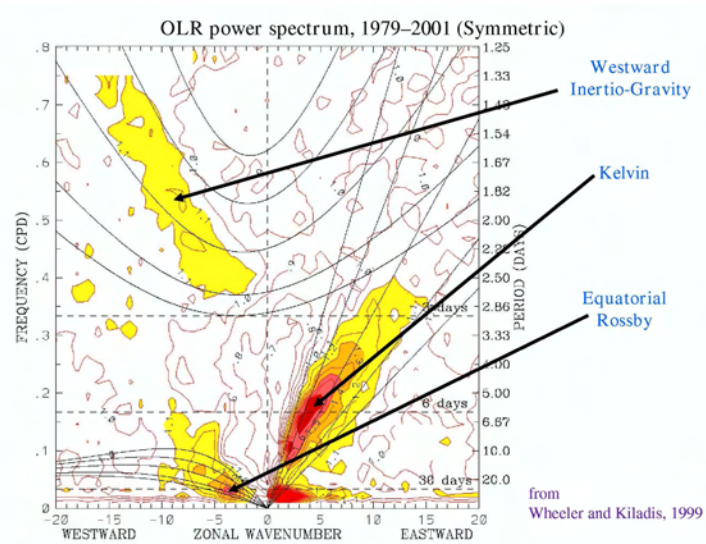


Figure 14: Symmetric components of the OLR. The spectral peaks correspond quite well to the dispersion relations of the equatorially trapped wave modes.

- orol. Atmos. Phys., 65 (1998), pp. 233–240.
- [2] J. W. DEARDORFF, *Convective velocity and temperature scales for the unstable planetary boundary layer and for rayleigh convection*, J. Atmos. Sci., 27 (1970), pp. 1211–1213.
 - [3] K. A. EMANUEL, *An air-sea interaction model of intraseasonal oscillations in the tropics*, J. Atmos. Sci., 44 (1987), pp. 2324–2340.
 - [4] ———, *Atmospheric Convection*, Oxford University Press US, 1994.
 - [5] ———, *A statistical analysis of hurricane intensity.*, Mon. Wea. Rev., 128 (2000), pp. 1139–1152.
 - [6] ———, *Tropical cyclones*, Annu. Rev. Earth Planet Sci., 31 (2000), pp. 75–104.
 - [7] N. O. RENNÓ AND E. R. WILLIAMS, *Quasi-lagrangian measurements in convective boundary layer plumes and their implications for the calculation of cape.*, Mon. Wea. Rev., 123 (1995), pp. 2733–2742.
 - [8] M. WHEELER AND G. N. KILADIS, *Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumbercfrequency domain.*, J. Atmos. Sci., 56 (1999), pp. 374–399.

Turbulent dynamos: Experiments, nonlinear saturation of the magnetic field and field reversals

C. Gissinger & C. Guervilly

November 17, 2008

1 Introduction

Magnetic fields are present at almost all scales in the universe, from the Earth (roughly 0.5 G) to the Galaxy (10^{-6} G). It is commonly believed that these magnetic fields are generated by dynamo action *i.e.* by the turbulent flow of an electrically conducting fluid [12]. Despite this space and time disorganized flow, the magnetic field shows in general a coherent part at the largest scales. The question arising from this observation is the role of the mean flow: Cowling first proposed that the coherent magnetic field could be due to coherent large scale velocity field and this problem is still an open question.

2 MHD equations and dimensionless parameters

The equations describing the evolution of a magnetohydrodynamical system are the equation of Navier-Stokes coupled to the induction equation:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla \pi + \nu \Delta \mathbf{v} + \mathbf{f} + \frac{1}{\mu \rho} (\mathbf{B} \cdot \nabla) \mathbf{B} , \quad (1)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \Delta \mathbf{B} . \quad (2)$$

where \mathbf{v} is the solenoidal velocity field, \mathbf{B} the solenoidal magnetic field, $\nabla \pi$ the pressure gradient in the fluid, ν the kinematic viscosity, \mathbf{f} a forcing term, μ the magnetic permeability, ρ the fluid density and η the magnetic permeability.

Dealing with the geodynamo, the minimal set of parameters for the outer core are :

- ρ : density of the fluid
- μ : magnetic permeability

- ν : kinematic diffusivity
- σ : conductivity
- R : radius of the outer core
- V : typical velocity
- Ω : rotation rate

The problem involves 7 independent parameters and 4 fundamental units (length L , time T , mass M and electric current A). Therefore there are 3 dimensionless parameters: the magnetic Reynolds number $Rm = \mu_0 \sigma V R$, the Reynolds number $Re = VR/\nu$ and the Rossby number $Ro = V/(R\Omega)$. One can also define the magnetic Prandtl number $Pm = \nu \mu_0 \sigma = Rm/Re$.

3 Numerical simulations, experiments and the universe

In 1995, the first direct numerical simulation of dynamo was obtained by Glatzmaier and Roberts [11]. The magnetic structure observed was very similar to the Earth's one but dimensionless parameters are up to ten orders of magnitude away from realistic values. For instance, Pm is of order 1 in these simulations whereas it is 10^{-5} in the Earth.

On the figure 1, we can observe that the experiments are closer to the natural objects in term of Pm than the numerical simulations. This situation is a strong motivation to carry out dynamo experiments. In such experiments, liquid sodium is used instead of liquid iron due to the higher electric conductivity. However taking into account rotation, the situation is more dramatical for experiments since there is more than ten order of magnitude for Rossby number between the Earth and experiments. Dynamo action can only occur for sufficient Rm , typically greater than 10. Since the Prandtl number is 10^{-5} in experiments, this yields very large Reynolds number $Re > 10^6$ and consequently very turbulent flow. The power needed to drive the flow is typically $P \sim \rho V^3 R^2 f(Re)$. Because of the low value of the viscosity, Re will be dropped in first approximation in the previous relation. This leads to

$$Rm \sim \mu_0 \sigma \left(\frac{PR}{\rho} \right)^{1/3} \quad (3)$$

Using liquid sodium, 100 kW are required to reach $Rm = 50$ with $R = 1m$. Moreover, to increase Rm by a factor 10, P should be increase by a factor 1000.

The first experimental dynamos were independently observed in 2001 in Riga [8], Latvia and Karlsruhe [19], Germany. Both experiments are based on analytical flows known to produce dynamo action for relatively low Rm .

The Riga experiment is based on the Ponomarenko flow with a well known dynamo threshold. It consists of a cylindrical pipe divided into three cylindrical shells. The external one is filled with liquid sodium at rest. The internal vessel contains an downward helical flow and the fluid recirculates upward in the intermediate shell (figure 2). The experimental

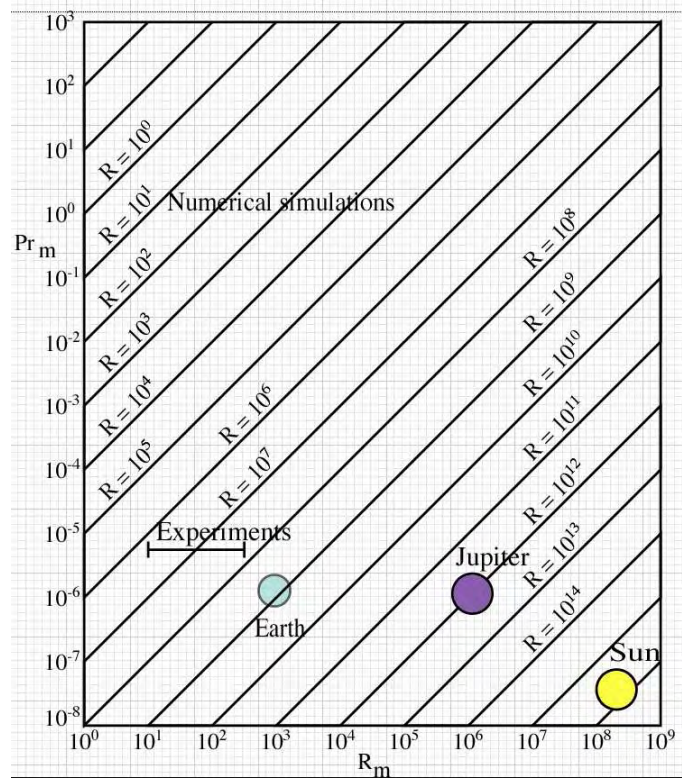


Figure 1: Parameter range for numerical simulations, experiments and natural systems.

threshold was found to be in good agreement with theoretical predictions. Moreover the Hopf bifurcation observed matched theory. The agreement between experiment and laminar kinematic dynamo theory is remarkable. However, using the non-linear theory the value of the saturated magnetic field is underestimated by a factor 10^6 .

The Karlsruhe experiment is based on the G.O. Roberts flow consisting in an array of whirling flow composed with azimuthal and axial flows (figure 2). The kinematic helicity defined by $(\nabla \times \mathbf{u}) \cdot \mathbf{u}$ is the same in all the pipes. The large scale magnetic field observed in this experiment involves an α – effect due to small scale flow. The threshold decreases when the large length scale increases but not if one decreases the small length scale.

In both experiments, the magnetic field is generated as if the mean flow were acting alone. This is probably due to the fact that the level of turbulence in these experiments is limited by the constraint of the pipes containing the flow.

The flow is very constrained in these two experiments and does not allow to study the effect of the turbulence on the dynamo action. In the situation of a bifurcation from a strongly turbulent flow, two questions arise:

- What is the effect of the turbulence on the dynamo threshold?
- What is the value of the saturated magnetic field above the dynamo threshold?

In order to study a turbulent flow, it can be useful to define the Reynolds decomposition

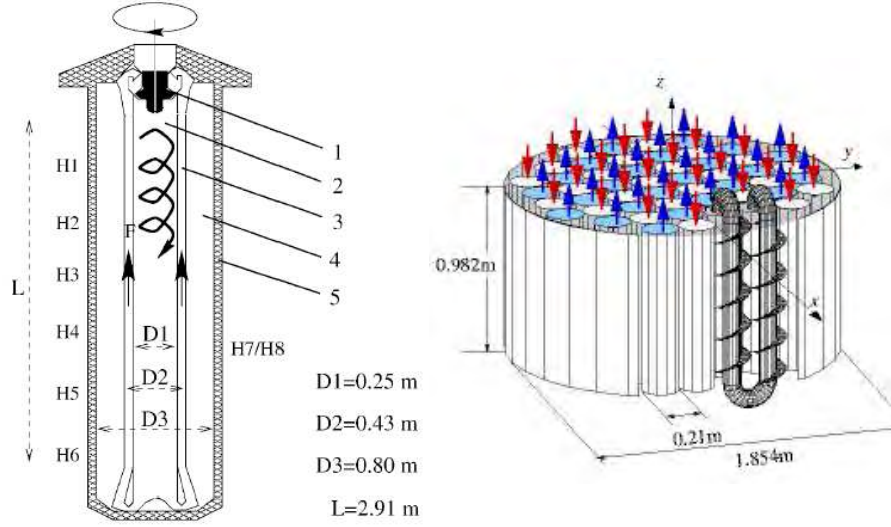


Figure 2: First experimental dynamos in constrained flows. Left: Riga experiment; Right: Karlsruhe experiment.

for the velocity field:

$$\mathbf{V}(\mathbf{r}, t) = \langle \mathbf{V} \rangle(\mathbf{r}) + \tilde{\mathbf{v}}(\mathbf{r}, t) \quad (4)$$

where $\langle \mathbf{V} \rangle$ represents the mean flow and $\tilde{\mathbf{v}}$ stands for the turbulent fluctuations. Then, the induction equation becomes

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\langle \mathbf{V} \rangle(\mathbf{r}) \times \mathbf{B}) + \nabla \times (\tilde{\mathbf{v}}(\mathbf{r}) \times \mathbf{B}) + \eta \Delta \mathbf{B} . \quad (5)$$

We see that the mean flow like the turbulent fluctuations can act as a source term for the magnetic field. There is two different approach for such problem: one can try to avoid large scale fluctuations using forced flow, like in Riga or Karlsruhe experiment. Another approach is to study the effect of velocity fluctuations on the dynamo instability which is the purpose of the VKS experiment.

4 The VKS experiment

4.1 Experimental set-up

The VKS (Von Karman Sodium) dynamo [13] is based on the so called Von Karman flow (figure 3). It consists of two toroidal cells in contrarotation driven by two impellers with 8 blades. Between two successive blades, the centrifugal flow is strongly expelled, creating a pumping of the flow near the axis. Both poloidal recirculation and azimuthal flow create a strongly turbulent velocity field with an important shear layer in the midplane due to the differential rotation. This flow is also known to produce a strong helical flow. The VKS flow is carried out in a cylindrical vessel, a geometry which, because of the fast rotation of

the Earth, appears to be not so far from a geodynamo situation. In this experiment the seed field is the magnetic field of the Earth.

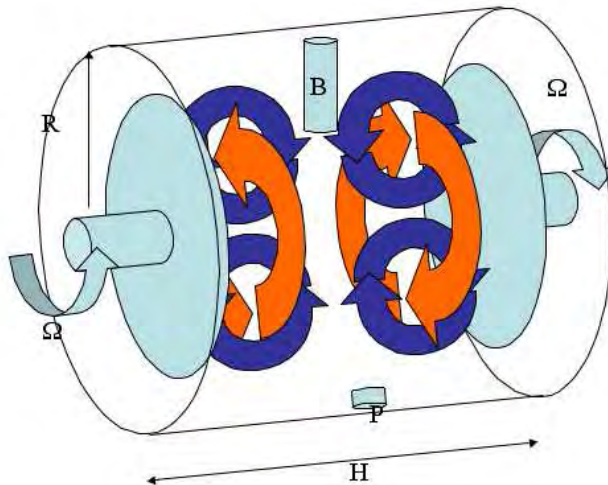


Figure 3: Sketch of the VKS flow.

There are other experimental attempts to observe self generation of magnetic field in a non-forced flow. The Madison experiment [7], in the University of Wisconsin, is based on Dudley & James flow and has been built to create a magnetic field due to this mean flow. Trials to observe dynamo in spherical couette flow, *i.e.* flow between two spherical shells in differential rotation are also made: the DTS (Derviche-Tourneur Sodium) experiment in Grenoble [14] and the Maryland experiment [18].

In the VKS experiment, 150-liters of liquid sodium are driven by the impellers with a power input of 300 kW. The temperature is monitored and 3 types of measurement are made: power, pressure and magnetic field. The last modification leading to the observation of a self-sustained magnetic field is the use of Iron discs with a strong magnetic permeability. This modification is known to reduce the threshold of the dynamo by changing the boundary conditions for the magnetic field [10]. The coercive field of the discs is close to 1 Gauss which is small compared to the magnitude of the magnetic field observed (100 G).

4.2 Results and interpretation

In exact contrarotation, first the generated magnetic field shows an exponential growth during the kinematic stage and then saturates at a strongly fluctuating state. The magnetic field is stationary and strongly axisymmetric. Generally the azimuthal part dominates the other components except near the axis, where the field is roughly an axial dipole. Figure 5 shows a typical evolution of the magnetic field of VKS.

The control parameter of the dynamo is the magnetic Reynolds number Rm . In the experiment, Rm can be varied by increasing the rotation rate of the propellers or by varying

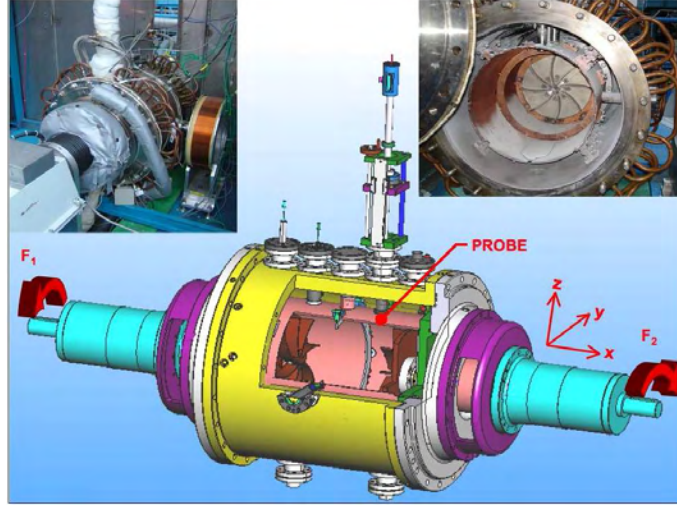


Figure 4: Sketch of the VKS experiment.

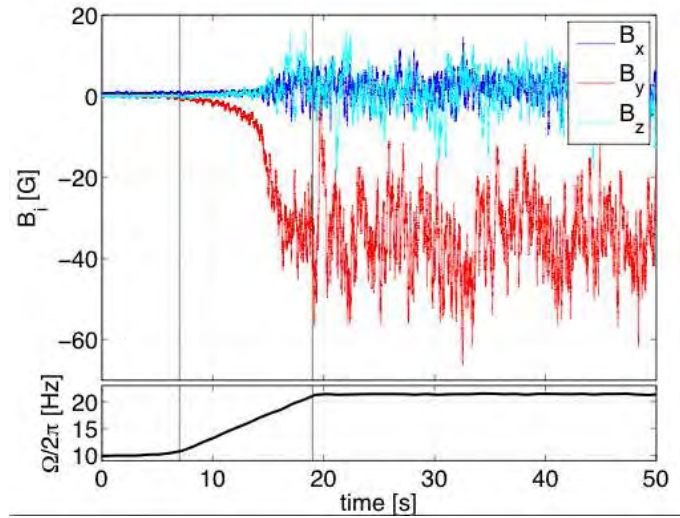


Figure 5: Time evolution of the magnetic field in the VKS experiment when the rotation rate of the discs is increased.

the temperature (and so the conductivity of sodium). The dynamo instability is a supercritical bifurcation with a threshold around $Rm_c = 30$. The saturated value of the magnetic field is in relatively good agreement with theoretical predictions (see lecture 2).

The generation of this magnetic field is understood as an $\alpha - \omega$ process [15]: due to the differential rotation between the two discs, the poloidal field is converted into a toroidal field by the so called ω -effect. In addition, the helicity of the flow will produce poloidal field from toroidal one: this α -effect is created by the flow near each disk. Between two successive blades of a disc, a strong centrifugal flow with helicity is created (figure 6). Thus, the VKS

dynamo is driven by the turbulent non-axisymmetric part of the flow. Indeed, numerical simulations based on the large scale mean flow alone always predicts an equatorial dipole breaking axisymmetry, in agreement with Cowling's theorem.

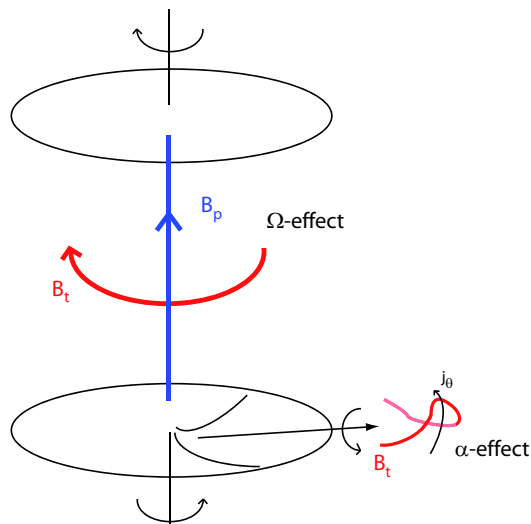


Figure 6: α and ω mechanisms in the VKS experiment.

Cowling's theorem states that an axisymmetric velocity field can not maintain an axisymmetric magnetic field. Because of the axisymmetrical nature of the forcing and the mean flow in the VKS experiment, one can be surprised by the observation of a strongly axisymmetric magnetic field. In fact, turbulent fluctuations have often been invoked to break the restriction imposed by Cowling's theorem. More surprising, even without such velocity fluctuations, it is possible to bypass the constraint of Cowling's theorem by a secondary bifurcation (figure 7). Indeed, the axisymmetric mean flow always generates a magnetic field breaking the axisymmetry, for instance an equatorial dipole in the case of the VKS experiment. The feedback of the Lorentz force immediately creates non-axisymmetric components in the velocity field, offering a way for the system to generate an axisymmetric magnetic field. This self-killing nature of the theorem can yield complex behavior like competition between equatorial and axial dipoles [9].

In the VKS experiment, the parameter space can be explored by imposing different rotation rates for each disk (non-exact counter-rotation). The figure 8 shows the different dynamics in the parameter space when the frequency of the propellers is modified. Three kinds of dynamics are observed: stationary, oscillatory and intermittent dynamos. Chaotic field reversals are also observed [2].

For example, decreasing the rotation rate of one propeller from 11 to 10 Hz and keeping the other one constant (28 Hz) yields a transition between a stationary regime to an oscillation. One can also notice subcritical bifurcation between oscillations and stationary field in the presence of bistability (figure 9). However, no transition occurs between the two

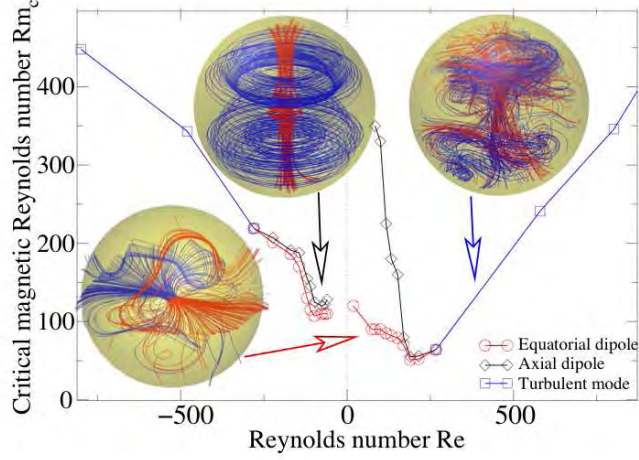


Figure 7: Numerical simulation showing how axial dipole can be generated using an axisymmetric forcing.

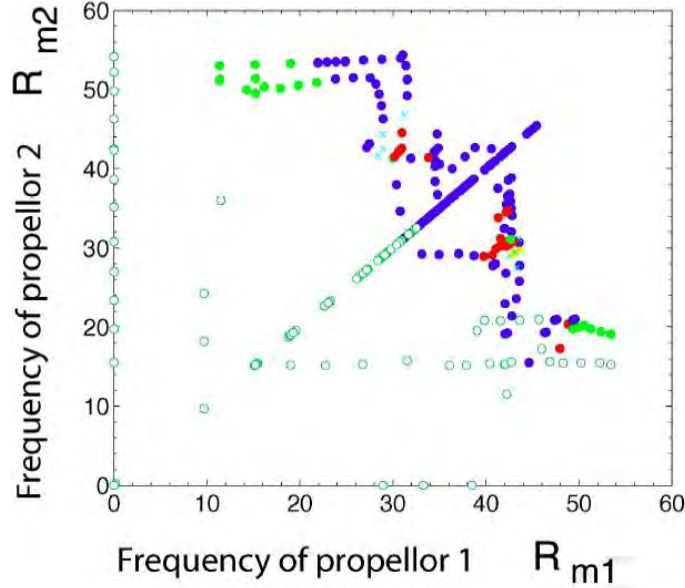


Figure 8: Different kinds of dynamics in the parameter space: open: no dynamo; closed dark: stationary dynamo; closed light: reversals and oscillatory dynamos

metastable regimes generated by the turbulent fluctuations on observable time scales.

The exploration of the parameter space shows that in non-exact counter-rotation, the observed relaxation oscillations bifurcate from fixed points located on the limit cycle as in the case of an excitable system. A very simple example of excitable system is a pendulum submitted to a constant torque. Indeed, when one apply a constant torque to a pendulum, the classical fixed points 0 (stable) and π (unstable) are respectively shifted to θ and $\pi - \theta$.

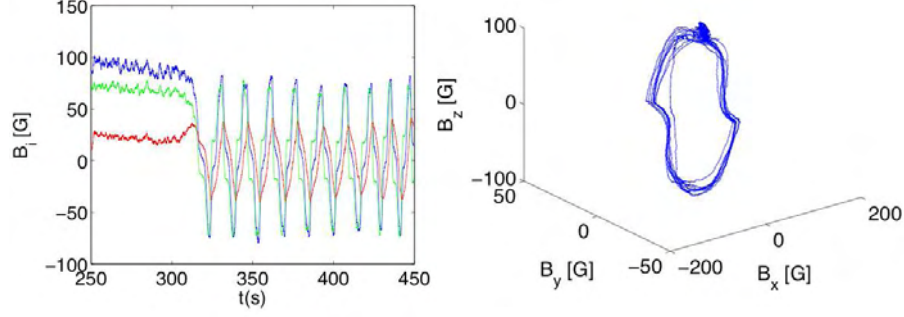


Figure 9: Transition from fixed point to a subcritical oscillatory dynamo. Left: time signal of the three components of the magnetic field. Right: Behavior in phase space.

When the torque is sufficiently strong, the stable and the unstable solutions collide and then bifurcate to a limit cycle. Near the relaxation oscillation, the dynamic of the field involves two different magnetic modes with close threshold: an axial dipole and a quadrupolar mode when propellers rotates at different frequencies. Experimentally we notice that a shift of the quadrupole along the rotation axis generates field reversals without variation of power.

The figure 10 shows a typical time signal of VKS experiment in the case of chaotic reversals. The dynamics seems very similar to the Earth and exhibits also some behavior called excursion, where the magnetic field temporarily decays but without reversing. Another interesting similarity with the Earth is the robustness of the reversals with respect to the turbulent fluctuations, especially the slow decay before a reversal followed by a fast recovery with a characteristic overshoot.

4.3 Model for VKS and geomagnetic reversals

This complicated behavior of the system can be well understood in the framework of a low-dimensional model, involving only the dipolar and quadrupolar modes [16]. The field can be decomposed as:

$$\mathbf{B} = d(t)\mathbf{D}(\mathbf{r}) + q(t)\mathbf{Q}(\mathbf{r}) \quad (6)$$

where \mathbf{D} and \mathbf{Q} represent the spatial structure of the dipolar and quadrupolar modes and $d(t)$ and $q(t)$ stand for the corresponding temporal evolution. In order to model the main feature of the magnetic field in a simple way, the study is restricted to the time evolution of the amplitudes d and q . The most general amplitude equations up to third order gives:

$$\dot{d} = \alpha_1 d + \beta_1 q + \gamma_{30} d^3 + \gamma_{21} d^2 q + \gamma_{12} d q^2 + \gamma_{03} q^3 \quad (7)$$

$$\dot{q} = \alpha_2 q + \beta_2 d + \eta_{03} q^3 + \eta_{21} d^2 q + \eta_{12} d q^2 + \eta_{30} d^3 \quad (8)$$

In exact counter-rotation, the problem presents a symmetry: a rotation of an angle π around a line in the equatorial plane. The two modes q and d can be classified into two different family depending on their behavior to this symmetry R_π . In the present case:

$$d \longrightarrow -d \quad (9)$$

$$q \longrightarrow q \quad (10)$$

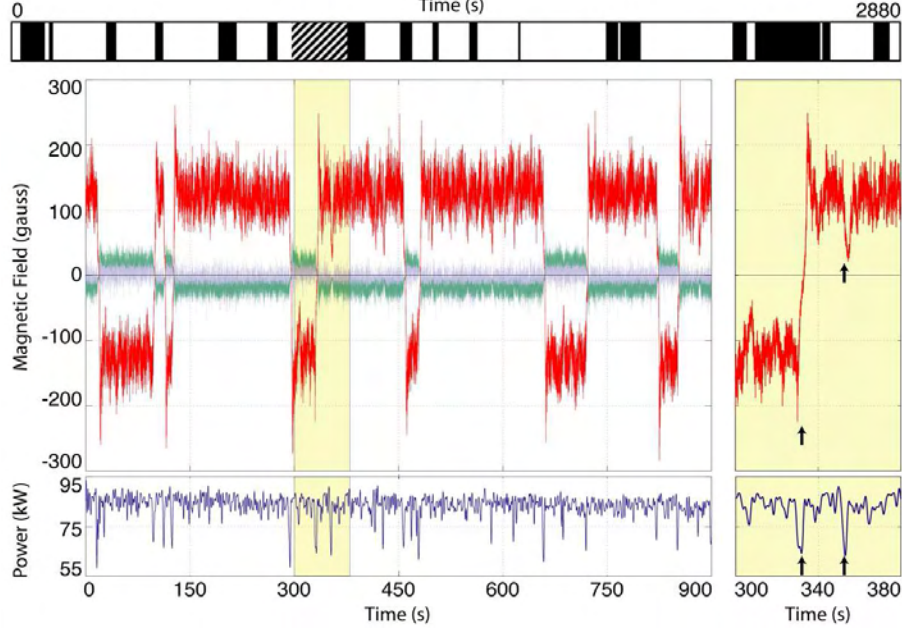


Figure 10: Time evolution of the magnetic field during reversal dynamics.

Under this symmetry, certain coefficients in equations 7 and 8 must vanish, giving

$$\dot{d} = \alpha_1 d + \gamma_{30} d^3 + \gamma_{12} d q^2 \quad (11)$$

$$\dot{q} = \alpha_2 q + \eta_{03} q^3 \eta_{21} d^2 q \quad (12)$$

Using the complex notation $A = d + iq = R e^{i\theta}$, we get the more compact form :

$$\dot{A} = \mu A + \nu \bar{A} + \lambda_1 A^3 + \lambda_2 A^2 \bar{A} + \lambda_3 A \bar{A}^2 + \lambda_4 \bar{A}^3 \quad (13)$$

where \bar{A} is the complex conjugate of A . In the general case, the coefficients are complex and depend on the experimental parameters. The system can be understood by looking at the linear part of equation 13. The time evolution of the angle θ is given by the imaginary part of equation 13

$$\dot{\theta} = \mu_i - \nu_r \sin 2\theta \quad (14)$$

We remark that when $\mu = 0$, the equations respect the symmetry R_π and describe the experiment in exact counter-rotation regime. In this case, the equation for θ leads to 2 fixed points, $\theta = 0$ and $\theta = \pi/2$, corresponding respectively to the stable axial dipole and the damped quadrupole (figure 11a). In the non-exact counter-rotation, μ is increased and the symmetry R_π is broken. Therefore dipole and quadrupole can interact. The equations for the modes become:

$$\dot{d} = (\mu_r + \nu_r) d - \mu_i q \quad (15)$$

$$\dot{q} = (\mu_r - \nu_r) q + \mu_i d \quad (16)$$

The two modes bifurcate to instability at very close thresholds when ν is sufficiently small. Therefore μ_i represents the symmetry breaking and corresponds to the difference of rotation rate between the two discs for the VKS experiment. When this symmetry is broken, the unstable solution evolves away from the purely quadrupolar mode and evolves an increasing amount of dipolar component. The two modes (stable dipole and unstable quadrupole) become closer. When the system reaches $\mu_i = \nu_c$, a bifurcation occurs: each stable solution collides with an unstable solution and disappears. This is a saddle node bifurcation that generates a limit cycle (thus an oscillatory magnetic field). This phenomenology is very similar to the one observed in the case of the pendulum with constant torque. However, a noteworthy difference occurs because of the additional symmetry $B \rightarrow -B$ yielding two other branches compared to the pendulum.

When $\mu_i \leq \nu_c$, the stable and unstable solutions are very close. When the system undergoes turbulent fluctuations, two different behaviors can arise:

- When the fluctuation is larger than $\nu_c - \mu_i$, the system jumps from the stable to the unstable fixed points and then bifurcates to the $-B$ stable solution. The system had undergone a reversal associated with an overshoot (figure 11b).
- When the fluctuation is less than $\nu_c - \mu_i$, the system can deviate from the stable solution without reaching the unstable one and comes back on the stable solution. This is very similar to geomagnetic events called excursions (figure 11c).

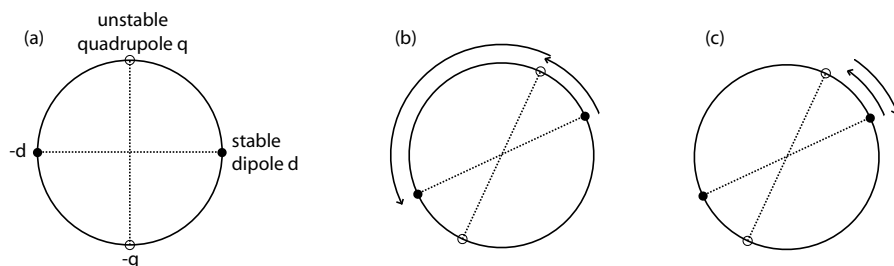


Figure 11: Three typical behaviors of the magnetic field. (a): Coexistence of two stationary solutions, stable and unstable when the symmetry R_π is respected. When R_π is broken, the interaction of dipolar and quadrupolar solutions can yield reversals (b) or simply excursions (c).

Using this simple model, the reversals of the VKS are understood as a consequence of the interaction of the axial dipole and the quadrupole. This interaction is related to the broken invariance of the flow under the rotation by π with respect to any axis in the mid-plane.

The magnetic field of the Earth also presents a competition between dipole and quadrupole. The symmetries involved in the case of the geodynamo are however different from the ones of the VKS experiment. In the core of the Earth, it is strongly believed that the flow is invariant under the mirror symmetry with respect to the equatorial plane. Under this

symmetry the dipole and the quadrupole can be classified into two different families. The previous model can thus naturally be applied to the case of the Earth. In this perspective, the reversals of the geomagnetic field could be caused by breaking the mirror symmetry of core dynamics.

5 Conclusion

- The VKS dynamo is not generated by the mean flow alone.
- The VKS experiment exhibits many different regimes in a small parameter range: stationary and oscillatory magnetic fields, reversals and excursions.
- A large scale dynamics of the field is observed: The results suggest that the system is governed by a few modes and this low dimensional dynamics is not smeared out by turbulent fluctuations.
- Using amplitude equation model, the reversals are understood to result from the competition between different modes (any external triggering mechanism are needed) and are due to a broken symmetry of the flow. A similar mechanism can be involved for planetary or stellar time dependent dynamos.

A Small and Large Re scaling of $\langle B^2 \rangle$

The flows creating the magnetic field of stars and galaxies involve huge kinetic, Re , and magnetic, Rm , Reynolds number. No laboratory experiments, nor direct numerical simulations are possible in the range of Re and Rm involved in astrophysical flows. It is thus interesting to try to guess scaling laws for the magnetic field using some simple hypotheses. This problem was first addressed by Batchelor for turbulent dynamos and is still an open question.

We consider here the minimum set of parameters: typical velocity V , length scale L , viscosity ν , magnetic permeability μ_0 and the fluid density ρ . We note that discarding global rotation makes our results certainly invalid for many astrophysical objects but not all of them. Rotation is indeed not assumed important for the galaxies which do not display a large scale coherent magnetic field [21],[20] and [4]. With 4 fundamental units, L (length), M (mass), T (time) and A (electric current), the problem involves 3 dimensionless parameters: $Rm = \mu_0 \sigma L V$, $Re = V L / \nu$ and $\langle B^2 \rangle / (\mu_0 \rho V^2)$. Some analytical calculations of threshold and saturation have been done for specific flow: for instance Childress and Soward for rotating *Rayleigh – Benard* convection [6] or Busse with a more complicated case [5]. However, these kind of calculations are always very difficult because the problem is not self-adjoint.

A.1 Laminar scaling

The total velocity field V_0 can be decomposed as follow:

$$V_0 = V_c + V_1 , \quad (17)$$

where V_c is the flow at the dynamo threshold and V_1 is some additional flow depending on the distance from the onset. For the laminar regime, the Stokes force $\rho\nu\Delta V_1$ balances the Lorentz force $\mathbf{J} \times \mathbf{B}$ in the saturated regime and we can easily evaluate the amplitude:

$$\frac{\rho\nu V_1}{l^2} \sim \frac{B^2}{\mu_0 l} . \quad (18)$$

Using $Rm = \mu_0\sigma l V_0$ and $Rm_c = \mu_0\sigma l V_c$, we are left with the expression:

$$B^2 \sim \frac{\rho\nu}{\sigma l^2} (Rm - Rm_c) , \quad (19)$$

for the mean magnetic energy density in the non-linear saturated regime. This is a weak field regime and it is not really relevant for the Earth which is very far from a laminar regime.

A.2 Large Re scaling

When Re is very large, the Lorentz force is now balanced by the inertial term and we get:

$$\frac{B^2}{\mu_0 l} \sim \rho \frac{V_0 V_1}{l} . \quad (20)$$

Expressing this equation in term of magnetic Reynolds number yields:

$$B^2 \sim \frac{\rho}{\mu_0(\sigma l)^2} (Rm - Rm_c) . \quad (21)$$

One can note that the ratio between the two scalings 19 and 21 is the magnetic Prandtl number Pm .

In Geophysics, the well known strong field regime is supposed to appear in a subcritical bifurcation from the weak field branch. It is simple to show that taking into account the rotation leads to this strong field regime. By balancing the Coriolis force and the Lorentz force, the scaling becomes thus:

$$B^2 \sim \frac{\rho\Omega}{\sigma} (Rm - Rm_c) . \quad (22)$$

This scaling shows why most of the direct numerical simulations (DNS) of geodynamo find a good agreement with the saturated value of magnetic field: all DNS generally manage to have the good ratio $\frac{\rho\Omega}{\sigma}$ despite Re or Ekman number being totally wrong.

B Turbulent Dynamos

As already stated, the problem of turbulent dynamos was proposed by Batchelor in 1950 [1]. In this paper, he looks at magnetic field generated by homogeneous isotropic turbulence with no mean helicity. Using a questionable analogy between the induction and the vorticity equations, he claimed that the dynamo threshold corresponds to $Pm = 1$, *i.e.* Rm_c proportional to Re , using our choice of dimensionless parameters. It is now often claimed

that Batchelor's criterion $Pm > 1$ for the growth of magnetic energy in turbulent flows is incorrect. It is however of interest to determine the minimal hypothesis for which Batchelor's predictions for dynamo onset is obtained using dimensional arguments. To wit, assume that the dynamo eigenmodes develop at small scales such that the threshold does not depend on the integral scale L . Then, discarding L in our set of parameters, dimensional analysis gives at once $Pm = Pm_c = \text{constant}$ for the dynamo threshold, *i.e.* Rm_c proportional to Re .

Another result of Batchelor concerns the saturation of the magnetic field. According to Batchelor's analogy between magnetic field and vorticity, the magnetic field should be generated mostly at the Kolmogorov scale, $l_K = LRe^{-3/4}$ where the velocity gradients are the strongest. He then assumed that saturation of the magnetic field takes place for $\langle B^2 \rangle / \mu_0$ proportional to $\rho v_K^2 K = \rho V^2 / \sqrt{Re}$ where v_K is the velocity increment at the Kolmogorov scale, $v_K^2 = \sqrt{\nu \epsilon}$. $\epsilon = V^3 / L$ is the power per unit mass, cascading from L to l_K in the Kolmogorov description of turbulence. ϵ being the power per unit mass available to feed the dynamo, it may be a wise choice to keep it, instead of V in our set of parameters, thus becoming $B, \rho, \epsilon, L, \nu, \mu_0$ and σ . Then, if we consider dynamo modes that do not depend on L , we obtain at once

$$\frac{B^2}{\mu_0} = \rho \sqrt{\nu \epsilon} h(Pm) = \frac{\rho V^2}{\sqrt{Re}} h(Pm) , \quad (23)$$

for saturation, where $h(Pm)$ is an arbitrary function of Pm . Close to dynamo threshold, $Pm \sim Pm_c$, we have $h(Pm) \propto Pm - Pm_c$ if the bifurcation is supercritical. This class of dynamos being small scale ones, it is not surprising that the inertial range of turbulence screens the magnetic field from the influence of integral size, thus L can be forgotten.

Some recent numerical simulations from Schekochihin [17] reported interesting behavior of the curve $Rm_c = f(Re)$. In these simulations of turbulent dynamo, one can observe that the threshold of the dynamo increases linearly with the Reynolds number on a large range of Re , according to Batchelor's theory. For sufficiently large Re , and using numerical hyperdiffusivity, the curve seems to saturate and to fall in the regime $Rm_c = \text{constant}$.

Others results from Bierman and Schluter [3] show results compatible with the scenario of Schekochihin for large Re . If Pm is small enough, the Kolmogorov and resistive scales are so far from each other that the magnetic field cannot feel the viscous dissipation. We can then drop ν in the analysis and we are left with $Rm_c = \text{constant}$. In addition, discarding ν in the limit $Pm \ll 1$ gives:

$$\frac{B^2}{\mu_0} = \rho V^2 g_0(Rm) , \quad (24)$$

where g_0 is an arbitrary function. Close to threshold, the rms velocity V is given by $\mu_0 \sigma V L \sim Rm_c$. In the case of a supercritical bifurcation $g_0(Rm) \propto Rm - Rm_c$ and we obtain

$$B^2 \propto \frac{\rho}{\mu_0 (\sigma L)^2} (Rm - Rm_c) . \quad (25)$$

Far from threshold, $Re \gg Rm \gg Rm_c$, one could assume that B no longer depends on σ provided that the magnetic field mostly grows at scales larger than l_σ . We then obtain

equipartition between magnetic and kinetic energy densities

$$\frac{B^2}{\mu_0} \propto \rho V^2, \quad (26)$$

as assumed by Biermann and Schluter.

C Evaluation of Ohmic dissipation

Ohmic losses due to currents generated by dynamo action give a lower bound to the power required to feed a dynamo. In order to evaluate them, it is crucial to know at which scales the magnetic field grows. Assuming that a dynamo is generated in the case $Pm \ll 1$, we want to give a possible guess for the power spectrum B^2 of the magnetic field as a function of the wave number k and the parameters ρ , ϵ , L , ν , μ_0 and σ . Far from threshold, $Re \gg Rm \gg Rm_c$, the dissipative lengths are such that $l_K \ll l_\sigma \ll L$. For k in the inertial range *i.e.* $kl_\sigma \ll 1 \ll kL$, we may use a Kolmogorov type argument and discard L , σ and ν . Then, only one dimensionless parameter is left and not too surprisingly, we get

$$|B|^2 \propto \mu_0 \rho \epsilon^{\frac{2}{3}} k^{-\frac{5}{3}}. \quad (27)$$

This is only one possibility among many others proposed for MHD turbulent spectra within the inertial range but it is the simplest. Integrating over k obviously gives the equipartition law for the magnetic energy. It is now interesting to evaluate Ohmic dissipation. Its dominant part comes from the current density at scale l_σ . We have

$$\frac{j^2}{\sigma} = \frac{1}{\sigma} \int |j|^2 dk \propto \frac{1}{\mu_0^2 \sigma} \int k^2 |B|^2 dk \propto \frac{\rho}{\mu_0 \sigma} \epsilon^{\frac{2}{3}} l_\sigma^{-\frac{4}{3}} \propto \rho \frac{V^3}{L}. \quad (28)$$

We thus find that Ohmic dissipation is proportional to the total available power which corresponds to some kind of optimum scaling law for Ohmic dissipation. Although, this does not give any indication that this regime is achieved, we note that the above scaling corresponds to the one found empirically from a set of numerical models. One can note that when the same derivation is applied for other scalings, different results are possible. For instance the k^{-1} spectrum yields a dissipation which can become greater than the input power. Thus the calculation of ohmic dissipation can invalidate the huge quantity of scaling present in the literature.

References

- [1] G. K. BATCHELOR, *On the Spontaneous Magnetic Field in a Conducting Liquid in Turbulent Motion*, Royal Society of London Proceedings Series A, 201 (1950), pp. 405–416.
- [2] BERHANU ET AL, *Magnetic field reversals in an experimental turbulent dynamo*, Europhys. Lett., 77 (2007), p. 59001.
- [3] L. BIERMANN AND A. SCHLÜTER, *Cosmic Radiation and Cosmic Magnetic Fields. II. Origin of Cosmic Magnetic Fields*, Physical Review, 82 (1951), pp. 863–868.

- [4] A. BRANDENBURG AND K. SUBRAMANIAN, *Astrophysical magnetic fields and nonlinear dynamo theory*, Phys. Rep., 417 (2005), pp. 1–4.
- [5] F. H. BUSSE, *Nonlinear interaction of magnetic field and convection.*, Journal of Fluid Mechanics, 71 (1975), pp. 193–206.
- [6] S. CHILDRESS AND A. M. SOWARD, *Convection-Driven Hydromagnetic Dynamo*, Physical Review Letters, 29 (1972), pp. 837–839.
- [7] FOREST ET AL, *Magnetohydrodynamics.*, 38 (2002), pp. 107–120.
- [8] GAILITIS ET AL, *Detection of a flow induced magnetic field eigenmode in the riga dynamo*, Phys. Rev. Lett., (2000).
- [9] C. GISSINGER, E. DORMY, AND S. FAUVE, *By-passing cowling’s theorem in axisymmetric fluid dynamos*, arxiv, (2008).
- [10] C. GISSINGER, A. ISKAKOV, E. DORMY, AND F. S., *Effect of magnetic boundary conditions on the dynamo threshold of von karman swirling flows*, Europhys. Lett., 82 (2008), p. 29001.
- [11] G. GLATZMAIER AND P. H. ROBERTS, *A three-dimensional self-consistent computer simulation of a geomagnetic field reversal*, Nature, 377 (1995), pp. 203–209.
- [12] H. MOFFATT, *Magnetic field generation in electrically conducting fluids*, Cambridge University Press, (1978).
- [13] MONCHAUX ET AL, *Generation of magnetic field by dynamo action in a turbulent flow of liquid sodium*, Phys. Rev. Lett., 98 (2007), p. 044502.
- [14] NATAF ET AL, *Geoph. Astroph. Fluids dyn.*, 100 (2006), pp. 281–298.
- [15] F. PETRELIS, N. MORDANT, AND S. FAUVE, *On the magnetic fields generated by experimental dynamos*, G. A. F. D., 101 (2007), p. 289.
- [16] PETRELIS ET AL, *geomagnetic reversals caused by breaking mirror symmetry of core dynamics*, arxiv, (2008).
- [17] A. A. SCHEKOCHIHIN, S. C. COWLEY, S. F. TAYLOR, G. W. HAMMETT, J. L. MARON, AND J. C. MCWILLIAMS, *Saturated State of the Nonlinear Small-Scale Dynamo*, Physical Review Letters, 92 (2004), pp. 084504–+.
- [18] SISAN ET AL, *Phys. Rev. Lett.*, 93 (2004), p. 114502.
- [19] STIEGLITZ ET AL, *Experimental demonstration of a homogeneous two-scale dynamo*, Phys. of Fluids, 98 (2001), p. 044502.
- [20] L. M. WIDROW, *Origin of galactic and extragalactic magnetic fields*, Reviews of Modern Physics, 74 (2002), pp. 775–823.
- [21] I. B. ZELDOVICH, A. A. RUZMAIKIN, AND D. D. SOKOLOV, eds., *Magnetic fields in astrophysics*, vol. 3, 1983.

Energetics of a Turbulent Ocean

Raffaele Ferrari

January 12, 2009

1 Energetics of a Turbulent Ocean

One of the earliest theoretical investigations of ocean circulation was by Count Rumford. He proposed that the meridional overturning circulation was driven by temperature gradients. The ocean cools at the poles and is heated in the tropics, so Rumford speculated that large scale convection was responsible for the ocean currents. This idea was the precursor of the thermohaline circulation, which postulates that the evaporation of water and the subsequent increase in salinity also helps drive the circulation. These theories compare the oceans to a heat engine whose energy is derived from solar radiation through some convective process.

In the 1800s James Croll noted that the currents in the Atlantic ocean had a tendency to be in the same direction as the prevailing winds. For example the trade winds blow westward across the mid-Atlantic and drive the Gulf Stream. Croll believed that the surface winds were responsible for mechanically driving the ocean currents, in contrast to convection. Although both Croll and Rumford used simple theories of fluid dynamics to develop their ideas, important qualitative features of their work are present in modern theories of ocean circulation.

Modern physical oceanography has developed a far more sophisticated picture of the physics of ocean circulation, and modern theories include the effects of phenomena on a wide range of length scales. Scientists are interested in understanding the forces governing the ocean circulation, and one way to do this is to derive energy constraints on the different processes in the ocean. The goal is to use these constraints to understand the forces that drive ocean circulation and the forces that maintain the density stratification. Using Sandstrom's theorem on horizontal convection it is possible to rule out convection as an important source of energy for the ocean circulation. Once this possibility has been ruled out we will consider the kinetic energy budget and study the effects of winds and tides on the ocean circulation [12].

1.1 Ocean Stratification

Figure 1 shows the density stratification in the Pacific Ocean[12]. The upper half kilometer of the ocean is highly stratified. Below this region there are still density changes but not to the same extent, and the density is much more uniform in the abyssal ocean. There is a region of highly dense water near the surface in the poles. This represents cold water

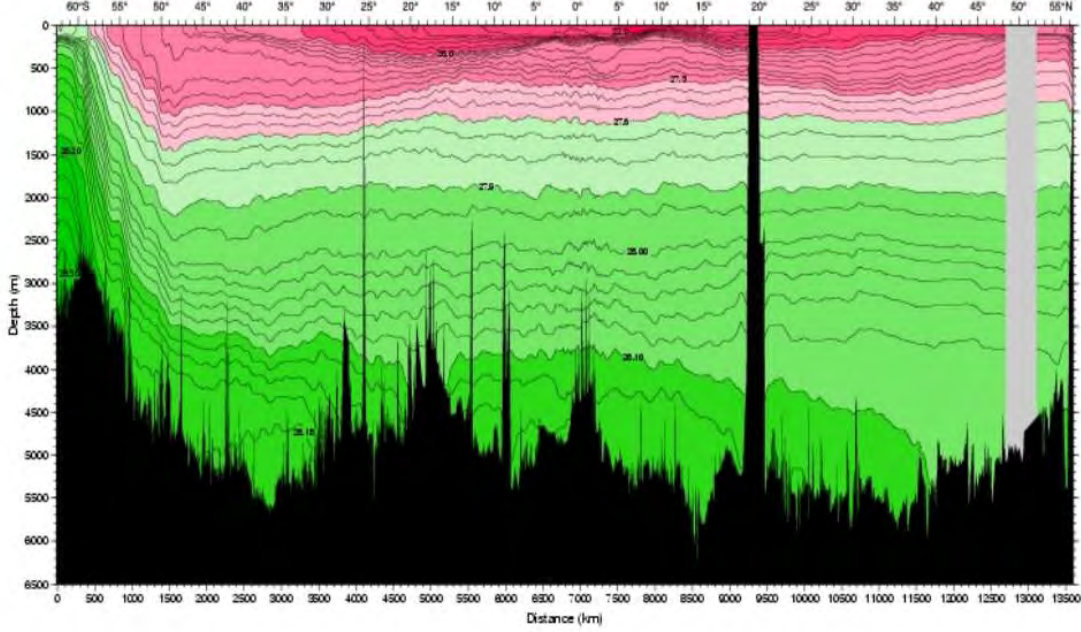


Figure 1: Potential density stratification in the Pacific Ocean. The level sets of pressure are closely packed near the surface where the ocean is highly stratified. In the deep ocean the pressure gradient weakens. Near south the pole the ocean is highly mixed, and there is a tongue of dense water sinking to the bottom.

from the surface, sinking and mixing with water from the abyss. There is a corresponding upwelling in other parts of the ocean that increases the potential energy. This upwelling requires an energy source to drive it.

1.2 The Oceanic Energy Budget

Throughout these notes we will use the Boussinesq approximation:

$$\frac{D\mathbf{u}}{Dt} + f\hat{z} \times \mathbf{u} = -\frac{1}{\rho_0}\nabla p + b\hat{z} + \nu\nabla^2\mathbf{u} - \nabla\Phi_{Tide} \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0 \quad (2)$$

$$\frac{D\theta}{Dt} = \kappa_\theta \nabla^2 \theta \quad (3)$$

$$\frac{DS}{Dt} = \kappa_S \nabla^2 S. \quad (4)$$

Here D is the advective derivative, θ is the potential temperature, and S is the salinity, b is the buoyancy, ν the molecular viscosity, and Φ_{Tide} the gravitational potential due to the tides. We assume that b is linear in θ and S :

$$b(\theta, S) = g\alpha\theta - g\beta S. \quad (5)$$

We take a linear combination of equations (3) and (4) to produce an expression $\frac{Db}{Dt}$ and eliminate θ and S from the equations:

$$\frac{Db}{Dt} = g\alpha \frac{D\theta}{Dt} - g\beta \frac{DS}{Dt} = \kappa \nabla^2 b. \quad (6)$$

The quantity $\kappa = \kappa_\theta + \kappa_S$. The total kinetic energy of the ocean is proportional to $\frac{1}{2} \int_{Ocean} \mathbf{u}^2 dV$. If take the dot product of the momentum equation by \mathbf{u} and integrate over the volume of the ocean we obtain the term by term kinetic energy budget of the ocean:

$$\frac{1}{2} \frac{\partial}{\partial t} \|\mathbf{u}\|^2 = - \int \frac{1}{2} |\mathbf{u}|^2 (\mathbf{u} - \mathbf{u}_s) \cdot \hat{n} dA - \int \left(\frac{wp}{\rho_0} - \frac{1}{2} \nu \nabla |\mathbf{u}|^2 \right) \cdot \hat{n} dA + \langle wb \rangle - \epsilon - \langle \mathbf{u} \cdot \nabla \Phi_{Tide} \rangle. \quad (7)$$

Here $\langle \cdot \rangle$ stands for the integral over the volume of the oceans, ϵ stands for $\langle \nu \|\nabla^2 \mathbf{u}\|^2 \rangle$, the viscous dissipation, and w is the velocity in the \hat{z} direction. If we ignore the effects of viscous dissipation at the boundaries and evaluate the dot products explicitly we find that:

$$\frac{1}{2} \frac{\partial}{\partial t} \|\mathbf{u}\|^2 = - \frac{1}{\rho_0} \int (wp - \tau \cdot b) dA + \langle wb \rangle - \epsilon - \langle \mathbf{u} \cdot \nabla \Phi \rangle. \quad (8)$$

The terms in this equation are the components of the kinetic energy budget. The $\langle wb \rangle$ represents the conversion of kinetic energy to potential energy. The dissipation ϵ represents the conversion of kinetic energy to internal energy. The other terms represent the work due to pressure, the tides, and friction stress at the boundary.

The potential energy is entirely due to gravity, and we define it to be $\langle -zb \rangle$. If we multiply the buoyancy equation by z we can determine the local potential energy change:

$$\frac{\partial}{\partial t} zb + \nabla \cdot (\mathbf{u}zb + z\kappa \nabla b) = wb - \kappa \frac{\partial b}{\partial z}. \quad (9)$$

The terms on the right come from moving z inside the divergence. If we integrate this equation and ignore the second term on the right we calculate the potential energy budget:

$$\frac{\partial}{\partial t} PE = \int z F_b dA - \langle wb \rangle + \left\langle \kappa \frac{\partial b}{\partial z} \right\rangle. \quad (10)$$

In this expression $F_b = \kappa \nabla b \cdot \hat{n}$. We recognize the familiar $\langle wb \rangle$ from the kinetic energy budget. There are also two other terms that represent the conversion of internal energy to kinetic energy.

The combination of the potential energy and kinetic energy budgets gives an energy budget for the entire ocean. Kinetic energy is generated by the winds and tides, and by conversion of potential energy. It is dissipated through viscosity into internal energy. The internal energy is generated by heat fluxes and dissipation and is converted into potential energy through buoyancy [12].

1.3 Convection

By eliminating the effects of the winds and the tides we can determine if the buoyancy fluxes are strong enough to drive the thermohaline circulation. We drop the tidal forces

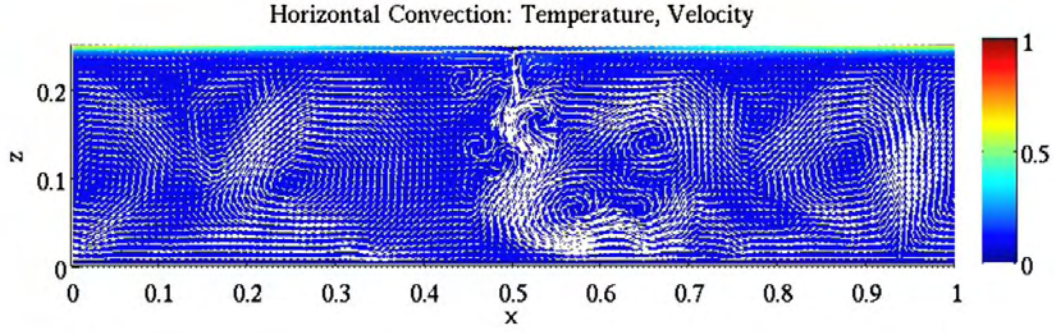


Figure 2: A numerical simulation of horizontal convection. There is a heat source on the edges and a heat sink in the center. Periodic boundary conditions are used. The arrows indicate the direction of the flow. The temperature is constant in most of the fluid and the heat flux is confined to the surface. There is no significant convective flow

and the boundary pressure and stress terms. The circulation in this ocean is driven only by convection. However the configuration of the ocean is different from standard convection because both the heating and cooling occur at the surface of the ocean.

Johan Sandstrom studied horizontal convection in systems where the heating was above and at the same height as the cooling[2]. This mimics the oceanic system where the heating and cooling occur at the same altitude but with a large horizontal displacement. Sandstrom argued that a heat engine does no work if the heating occurs at a higher or equal pressure than the cooling. This argument translates to the ocean, where the term $-\int zF_b dA$, the work done by buoyancy, is either 0 or negative.

The application of Sandstrom's conjecture to the ocean has been validated by other numerical and theoretical work. If we take the limit $\kappa \mapsto 0$ with the ratio $\frac{\nu}{\kappa}$ constant the ocean becomes stagnant and homogeneous. The energy conversion between internal energy and potential energy ceases, and vice versa. Lab experiments and numerical simulations by Rossby show explicitly that the overturning stream function vanishes. Another example is Figure 2, which is a two dimensional numerical simulation of horizontal convection. All of the temperature gradients are confined to the surface and there is no convective flow.

There are actual numerical estimates for the strength of convection in the ocean. For example Faller and Huang[5] have estimated the order of magnitude of energy conversion between internal and potential energy to be:

$$-\rho_0 g^{-1} \int z \bar{F}_b dA \approx \pm .01 TW \quad (11)$$

$$\rho_0 g^{-1} \langle \kappa \bar{\partial}_z b \rangle \approx \pm .001 TW. \quad (12)$$

Similarly $\langle wb \rangle$ is zero because of Sandstrom's theorem. Therefore there is no appreciable energy flow from potential energy to kinetic or from internal energy to potential. This means that the circulation of the ocean is driven by the winds and tides and that this energy is converted into internal energy through viscous dissipation. This internal energy is returned to the atmosphere through heat fluxes.

1.4 Kinetic Energy Sources

If the ocean is not driven by heat it must be driven mechanically. The source of these forces are surface winds and tides. The surface work term can be decomposed into three important components, the geostrophic flow, ageostrophic flow, and surface waves:

$$w_{wind} = \int \tau \cdot \mathbf{u}_g dA + \int \tau \cdot \mathbf{u}_{ag} dA + \int (\tau \cdot \mathbf{u}_w - pw) dA. \quad (13)$$

The geostrophic flow comes from balancing the Coriolis force with the atmospheric pressure. The wind is a predominantly geostrophic flow, and the geostrophic flow is directed along the level sets of the pressure. The ageostrophic flow is the actual flow minus the geostrophic flow. Scientists have tried to estimate the energy transfer due to each of these terms. The geostrophic flow contributes approximately $.8TW$, the ageostrophic flow $3TW$, and the surface waves $60TW$. The actual kinetic energy transfer to the circulation of the ocean below the turbulent boundary layer at the surface is more difficult to determine. It is believed that the surface waves are dissipated in the form of turbulence and that they cannot create kinetic energy below the turbulent boundary layer. Likewise it is believed that only $.5TW$ of energy from the ageostrophic flow penetrates the boundary layer. It is thought that all of the energy stress from the geostrophic flow is transferred to kinetic energy in the surface boundary layer[1].

The other major source of mechanical energy is the tides. The overall power due to tidal forces is roughly $3.5TW$. Most of this, $2.5TW$ is done in shallow seas against the continental shelves. About $1TW$ is done against the abyssal topography. It is believed that only $1TW$ of this is converted into useful kinetic energy in the ocean. This means that the global transfer of power from the winds and tides into ocean kinetic energy is roughly $2.3TW$. Therefore we would expect the dissipation of kinetic energy into internal energy to be $2.3TW$ in order to balance the inputs[3].

2 The Energy Cascade

The ocean contains energy on a wide range of length scales and frequencies. The oceanic kinetic energy is dominated by geostrophic turbulence, i.e. the energy in the mesoscale eddy field. Figure 3 shows two kinetic energy density spectra from moorings in the North Atlantic and the Southern Ocean. Note that the spectra are predominantly red. On average about 90% of the kinetic energy in the ocean is in subinertial frequencies, i.e. timescales longer than the local pendulum day $2\pi/f$. On smaller scales¹ the kinetic energy is dominated by internal waves in the inertial peak and the semi-diurnal lunar tide peak. Only a very small fraction of the kinetic energy lies in the small scale 3 dimensional turbulence. However, the latter is an important pathway to the final dissipation of kinetic energy. The bulk of the potential energy on the other hand is in the planetary scale ocean stratification.

¹We use frequencies and wave numbers somewhat interchangeable in this discussion. This can in general not strictly be justified, though it should be acceptable for a rough qualitative discussion as given here. Spectral energy estimates in wavenumber space are generally much harder to obtain from observational data.

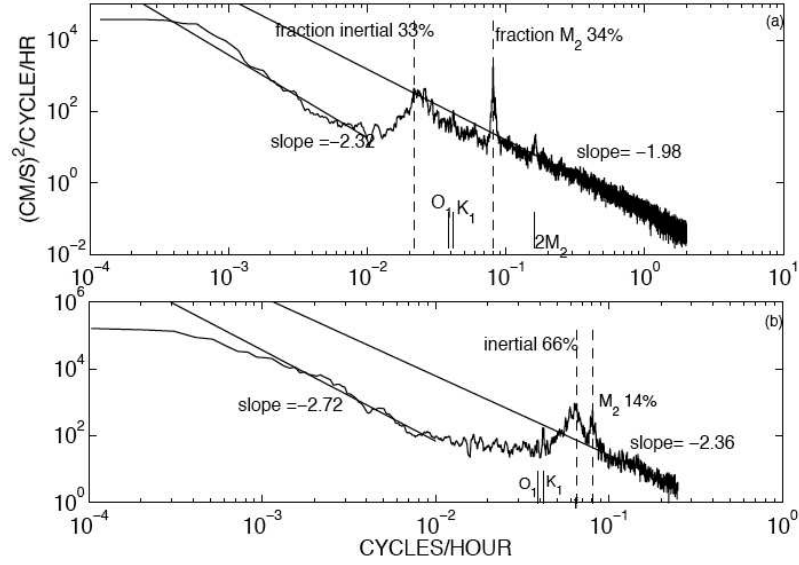


Figure 3: (a) Kinetic energy estimate for an instrument in the western North Atlantic at about 15° N at 500m. This record was described in [6]. (b) Power density spectral estimate from a record at 1000m at 50.7° S, 143° W, south of Tasmania in the Southern Ocean [8]. The inertial and principal lunar, semi-diurnal M_2 , tidal peaks are marked, along with the percentage of kinetic energy in them lying between f and the highest frequency estimate. Least-squares power law fits for periods between 10 and 2 hours and for periods lying between 100 and 1000 hours are shown. The approximate percentage of energy of the internal wave band lying in the inertial peak and the M_2 peak are noted.

2.1 Pathways of Kinetic Energy: From Forcing to Dissipation

The generation of kinetic energy below the surface mixed layer acts primarily at planetary scales. Most of the energy is generated by the large-scale wind field and by the tides. Dissipation on the other hand occurs at scales on the order of $O(1cm)$ or $O(1s)$ or less. Thus energy has to be transferred across the spectrum from the largest scales to the smallest scales. How does this happen? It is well understood that a forward energy cascade acts in 3-dimensional isotropic turbulence. I.e. turbulent kinetic energy is transported to smaller scales until it gets finally dissipated. However, the situation in the ocean is much more complicated. There are (at least) two main difference. Due to the presence of stratification and rotation larger scale turbulence in the ocean is obviously not isotropic. Moreover, exchanges between kinetic and potential energy play an important role.

2.1.1 Conversion Between Kinetic Energy and Potential Energy

As argued before, the total conversion between kinetic and potential energy, given by the conversion term $\overline{\langle wb \rangle}$ approximately has to vanish. However, significant conversions can occur locally or over certain spatial and temporal scales. For the following discussion we want to decompose variables into three different scales. The buoyancy field for example is given by

$$b = \bar{b} + b_e + b_t.$$

The first term on the right hand side describes the buoyancy field associated with the mean circulation, i.e. planetary scales. b_e denotes the contribution of the geostrophic eddy field and b_t describes small scale 3D turbulence. Assuming that this scale separation can strictly be made, the $PE \rightarrow KE$ conversion term can be decomposed into

$$\overline{\langle wb \rangle} = \overline{\langle w_t b_t \rangle} + \overline{\langle w_e b_e \rangle} + \overline{\langle \bar{w} \bar{b} \rangle} \quad (14)$$

The mean $PE \rightarrow KE$ conversion $\overline{\langle \bar{w} \bar{b} \rangle}$ is dominated by the generation of potential energy by winds² and by the abyssal overturning circulation which releases potential energy. It can be written as

$$\begin{aligned} \overline{\langle \bar{w} \bar{b} \rangle} &= \overline{\langle \bar{w}_{Ekman} \bar{b} \rangle} + \overline{\langle \bar{w}_{ag} \bar{b} \rangle} \\ &= - \iint (\tau_{wind} \cdot \bar{\mathbf{u}} - \tau_{bottom} \cdot \bar{\mathbf{u}}) dA + \overline{\langle \bar{w}_{ag} \bar{b} \rangle} \\ &\approx -0.8 \pm 0.1 \text{TW} + 0.1 \pm 0.1 \text{TW} + 1.0 \pm 0.5 \text{TW} \end{aligned} \quad (15)$$

The first terms on the RHS describe the production of potential energy by the Ekman pumping/suction due to the global wind field and a probably small sink of PE by Ekman pumping/suction due to bottom friction. Note that the error estimate for the latter is on the order of 100%. The last term describes the production of kinetic energy by subsidence in the (ageostrophic) abyssal overturning circulation. The two contributions are of opposite sign and similar within the error estimates.

²The large scale wind field is often regarded to produce potential energy. This happens due to Ekman pumping which, however, is indeed a conversion from KE to PE. More accurately, the wind field produces kinetic energy of which the major part is instantaneously converted into potential energy by geostrophic adjustment.

The mean currents are associated with strong buoyancy gradients and shear flows which become unstable and generate eddies. This process releases potential energy into geostrophic eddy kinetic energy. Since this is associated with a spindown of the large scale circulation, the second term in (14) can be expressed as an effective eddy stress τ_{eddy} acting on the large scale circulation:

$$\overline{\langle w_e b_e \rangle} = \iiint \mathbf{u} \cdot \partial_z \tau_{eddy} dV \approx 0.5 \pm 0.1 \text{TW} \quad (16)$$

Finally turbulent dissipation of kinetic energy results in an increase in potential energy. Osborn (1980) [7] argued that, in a stably stratified fluid, most of the turbulent kinetic energy is dissipated by viscous friction, while a fraction Γ is used to vertically mix the fluid and thus increase the potential energy.

$$|\overline{\langle w_t b_t \rangle}| = |\Gamma \overline{\langle \nu ||\nabla \mathbf{u}||^2 \rangle}| \leq \Gamma \epsilon \approx 0.5 \text{TW},$$

where ν is the viscosity and ϵ is the total dissipation of kinetic energy in the deep ocean, which has to equal the total input of kinetic energy into the latter. The mixing efficiency Γ in the stratified ocean was estimated to be around 20%. Estimates from direct observations [11] suggest

$$\overline{\langle w_t b_t \rangle} \approx -0.4 \pm 0.1 \text{TW} \quad (17)$$

Summing up the estimates of equations (15), (16) and (17) we find that the estimated total conversion is about zero. While the sign of the PE \rightarrow KE conversion by the large scale circulation is not clear, the meso-scale eddy field releases PE while turbulent mixing increases PE. The PE balance in the upper ocean (here the upper 1000-2000 meters) is dominated by the potential energy released by eddies which is maintained by the large-scale winds. In the abyss the release of potential energy by the deep overturning circulation is maintained by mixing and by the large scale wind forcing.

2.1.2 The Energy Cascade and Dissipation of Kinetic Energy

In the previous section we argued that kinetic energy is produced from potential energy at the geostrophic eddy scale while it is removed by small scale 3-dimensional turbulence. In this section we try to address the question how the energy is transported to smaller scales where it can finally be dissipated by turbulence.

Kinetic energy in inertial and tidal waves can be transferred to small scales through wave-wave interactions in the internal wave field and can finally be dissipated through breaking of internal waves. The estimated energy dissipated by internal wave breaking is about 1.5 TW.

It can be shown that the viscous dissipation of kinetic energy ϵ is proportional to the enstrophy $Z \equiv \frac{1}{2}|\zeta|^2 = \frac{1}{2}|\nabla \times \mathbf{u}|^2$:

$$\frac{\partial E}{\partial t} + \nabla \cdot (\mathbf{u}p + \mathbf{u}E - \nu \nabla E) = -2\nu Z + F. \quad (18)$$

In three dimensional turbulence, enstrophy can be generated by vortex stretching

$$\frac{\partial Z}{\partial t} + \nabla \cdot (\mathbf{u}Z + \nu \nabla Z) = \zeta_i S_{ij} \zeta_j - \nu ||\nabla \zeta||^2, \quad (19)$$

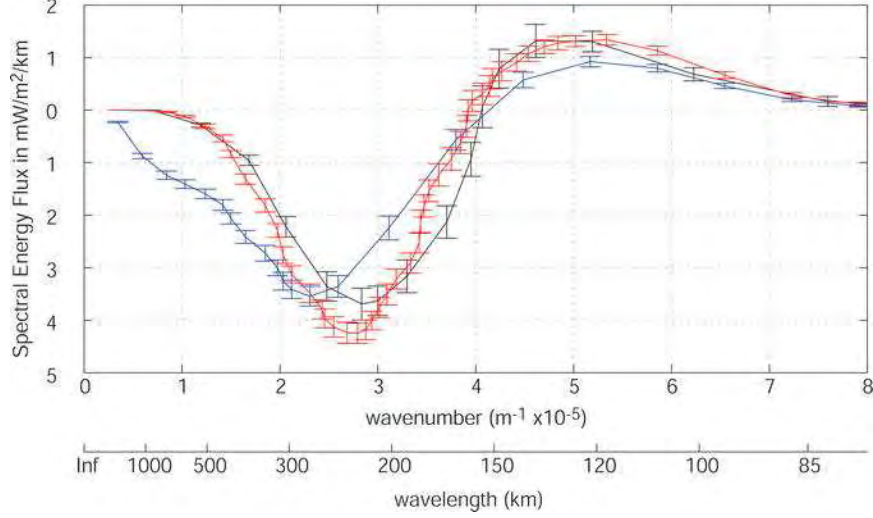


Figure 4: Time mean, spectral kinetic energy flux vs total wavenumber in a homogeneous ACC region (rectangles centered at 57S, 120W): black curve using sea surface height (SSH) on a 32×32 grid, red curve using SSH on a 64×64 grid, blue curve using velocity on a 64×64 grid. Positive slope reveals a source of energy. The larger negative lobe reveals a net inverse cascade to lower wavenumber. Error bars represent standard error. Taken from [10]

where $S_{ij} = \frac{1}{2}(\partial_{x_i} u_j - \partial_{x_j} u_i)$. The first term on the right hand side describes the change of enstrophy due to vortex stretching.

In two dimensional geostrophic turbulence (i.e. the mesoscale eddy field), however, enstrophy is conserved in the limit that the viscosity goes to zero. We find

$$\frac{\partial Z}{\partial t} + \nabla \cdot (\mathbf{u}Z + \nu \nabla Z) = -\nu \|\nabla \zeta\|^2, \quad (20)$$

where $\zeta = \partial_x v - \partial_y u$ is now a scalar. Since enstrophy is conserved in the limit $\nu \rightarrow 0$ the viscous energy dissipation has to vanish in geostrophic turbulence. In 3D turbulence on the other hand enstrophy is generated through vortex stretching and thus viscous energy dissipation remains finite as $\nu \rightarrow 0$ (since $Z \rightarrow \infty$).

Indeed it can be shown that in 3D isotropic turbulence, energy is transported to smaller scales until it is finally dissipated. The magnitude of the viscosity only determines the scale at which the energy is finally dissipated. In geostrophic turbulence on the other hand energy is transported to larger scales and is thus not dissipated. This "inverse cascade" is generally only limited by the planetary scale or the " β -effect" (see [9] for details). Figure 4 shows an observational estimate of the spectral energy flux in wavenumber space in a region in the Antarctic Circumpolar Current. We can see a weak forward cascade for wavelength of less than about 150km and a dominating inverse cascade on larger scales. The bulk of the ocean kinetic energy resides in the geostrophic eddy field and is thus transported to larger horizontal scales and generally also larger vertical scales, i.e. the eddy field tends to become barotropic. So where does the energy go from here?

2.2 Geostrophic Kinetic Energy Dissipation

Once the Eddies become barotropic, they will lose energy through bottom boundary layer drag, i.e. kinetic energy is dissipated in the bottom Ekman layer. The estimated energy loss through this process is

$$\epsilon_b = \iint \rho C_d |\mathbf{u}|^3 dA \approx 0.2 - 0.8 \text{ TW}. \quad (21)$$

Another important mechanism for dissipation of geostrophic KE arises from geostrophic flows over topography. Flow with velocity U impinging over topographic features with wavenumbers between f/U and N/U can cause radiation of internal waves which tend to interact and break within about 1 km above the topography. Observations indeed show strongly enhanced mixing over topography which can be attributed to this process. This is likely to be an important contributor to abyssal mixing. The estimated dissipation of KE due to this process is

$$\epsilon_t = \iiint \rho \nu |\nabla \mathbf{u}|^2 dV \approx 0.2 - 0.4 \text{ TW}, \quad (22)$$

where \mathbf{u} here denotes the velocity field due to topographically generated internal waves. In the surface boundary layer, particularly in the regions of the boundary currents, the geostrophic velocity field can form very sharp fronts due to frontogenesis. At large Richardson numbers, i.e. weak stratification, these fronts can go unstable and form small scale turbulence which is dissipated in a forward energy cascade. Since this "loss of balance" is confined to the surface boundary layer, the energy loss due to this pathway does not contribute to the mixing necessary to maintain the deep overturning circulation.

2.3 The Oceanic Energy Cycle

As an attempt to summarize the pathways of energy in the ocean, we found that in the upper ocean, the circulation is powered by the winds. Eddies are generated by baroclinic instabilities of the large scale circulation. Eddy-eddy interaction then transfers energy into barotropic motion following the geostrophic inverse energy cascade. Some of the energy in the eddy field is lost by loss of balance in the surface layer, where sharp fronts can evolve that finally become unstable. Another part is lost to bottom friction. Moreover, the geostrophic flow impinging over topography can generate internal waves which interact and break. This results in abyssal mixing that powers the overturning circulation.

A much more complete picture of the oceanic energy cycle is sketched in figure 5. For further details, the reader is referred to [4] and [12].

References

- [1] M. ALFORD, *Improved global maps and 54-year history of wind-work on ocean inertial motions*, Geophysical Research Letters, 30(8) (2003).
- [2] A. DEFANT, *Physical Oceanography*, Pergamon Press, 1961.

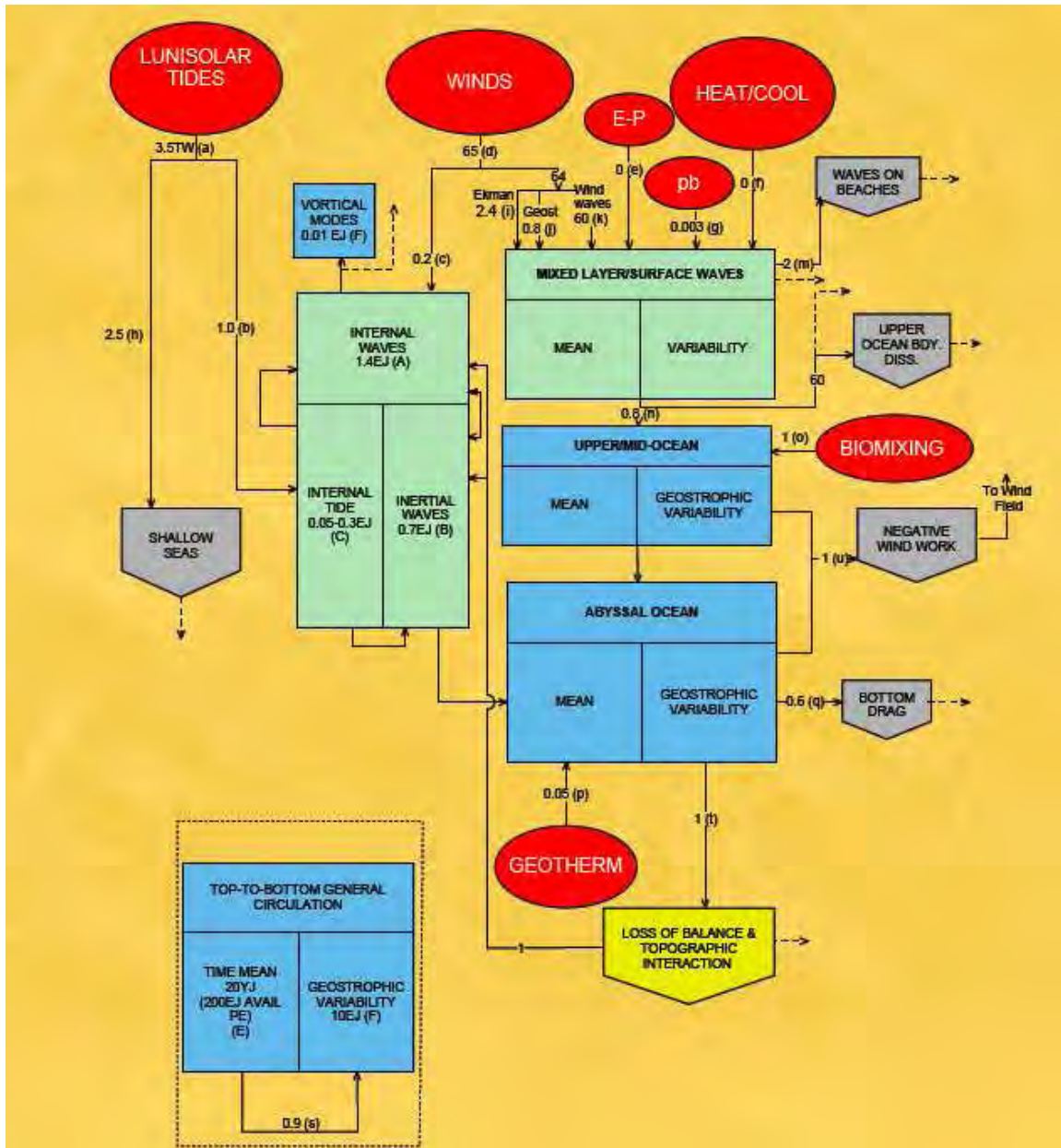


Figure 5: Schematic of the estimated pathways and reservoirs of energy in the ocean. All fluxes are given in TW. From [4].

- [3] G. EGBERT AND R. RAY, *Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data*, Nature, 405 (2000), pp. 775–778.
- [4] R. FERRARI AND C. WUNSCH, *Ocean circulation kinetic energy – reservoirs, sources and sinks*. In preparation for Annu. Rev. Fluid Mech., 2008.
- [5] R. HUANG AND W. WANG, *Gravitational potential energy sinks in the ocean*.
- [6] L. L. FU, T. KEFFER, P. NIILER, AND C. WUNSCH, *Observations of mesoscale variability in the western North Atlantic*, J. Mar. Res., 40 (1982), pp. 809–848.
- [7] T. R. OSBORN, *Estimates of the local rate of vertical diffusion from dissipation measurements*, J. Phys. Oceanogr., 10 (1980), pp. 83–89.
- [8] H. E. PHILLIPS AND S. R. RINTOUL, *Eddy variability and energetics from direct current measurements in the Antarctic Circumpolar Current south of Australia.*, J. Phys. Oceanogr., 30 (2000), pp. 3050–3076.
- [9] P. B. RHINES, *Geostrophic turbulence*, 11 (1979), pp. 401–441.
- [10] R. G. SCOTT AND F. WANG, *Direct evidence of an inverse kinetic energy cascade from satellite altimetry*, J. Phys. Oceanogr., 35 (2005), pp. 1650–1666.
- [11] L. ST. LAURENT AND H. SIMMONS, *Estimates of power consumed by mixing in the ocean interior*, J. Climate, 6 (2006), pp. 4877–4890.
- [12] C. WUNSCH AND R. FERRARI, *Vertical mixing, energy, and the general circulation of the oceans*, Annu. Rev. Fluid Mech., 36 (2004), pp. 281–314.

Ocean mixing, Internal tides and Tidal power

Chris Garrett

19 June 2008

1 Ocean mixing and Internal tides

1.1 Introduction

Mixing of the oceans ultimately occurs through molecular diffusion, though it is driven by fluctuations in the velocity, temperature and salinity on length scales down to 1 mm that cannot be resolved in large 3D numerical models which typically have a resolution of the order of $10 \text{ km} \times 10 \text{ km} \times 100 \text{ m}$. It will be many, many, years before resolution can be improved to resolve the small scale turbulence which causes mixing; as a result accurate parameterisations of the small scale processes are necessary.

Sub-grid scale mixing in the ocean may be likened to the influence of clouds or ice crystals in the atmosphere - they are unresolved in any large scale model, and yet have a very important effect on the overall dynamics. The mixing processes in the ocean are less obvious without injecting artificial dyes but likewise play an important role in the global oceanic circulation.

The general framework to parameterise mixing is usually to represent the eddy flux of a scalar C as a tensor multiplying the gradient of the mean \bar{C} ;

$$\overline{u'_i C'} = -T_{ij} \frac{\partial \bar{C}}{\partial x_j}, \quad (1)$$

where u'_i and C' are the small scale fluctuations of the velocity components and the scalar. The antisymmetric part of T_{ij} has an associated 'skew flux' $\mathbf{D} \times \nabla \bar{C}$ which is parallel to surfaces of constant \bar{C} . The symmetric part of T_{ij} may be diagonalised and thus represents diffusion parallel to some principal axes. The skew flux and large isopycnal diffusivities are likely to be a consequence of stirring by mesoscale eddies; the flux of more interest for ocean mixing is that *across* isopycnals, i.e. the diapycnal flux, and this is parameterised in terms of an eddy diffusivity K_v . Some possible causes for concern with this approach were noted; there are an implicit assumptions that there is some local 'mixing length' and that there is a spectral gap between the mean and fluctuations; in reality this distinction may be rather blurred. Nevertheless, quantifying mixing with a diffusivity is the common method used in numerical models; the question is, how should it vary in space and time, and what causes these variations?

Knowledge of diapycnal mixing comes from

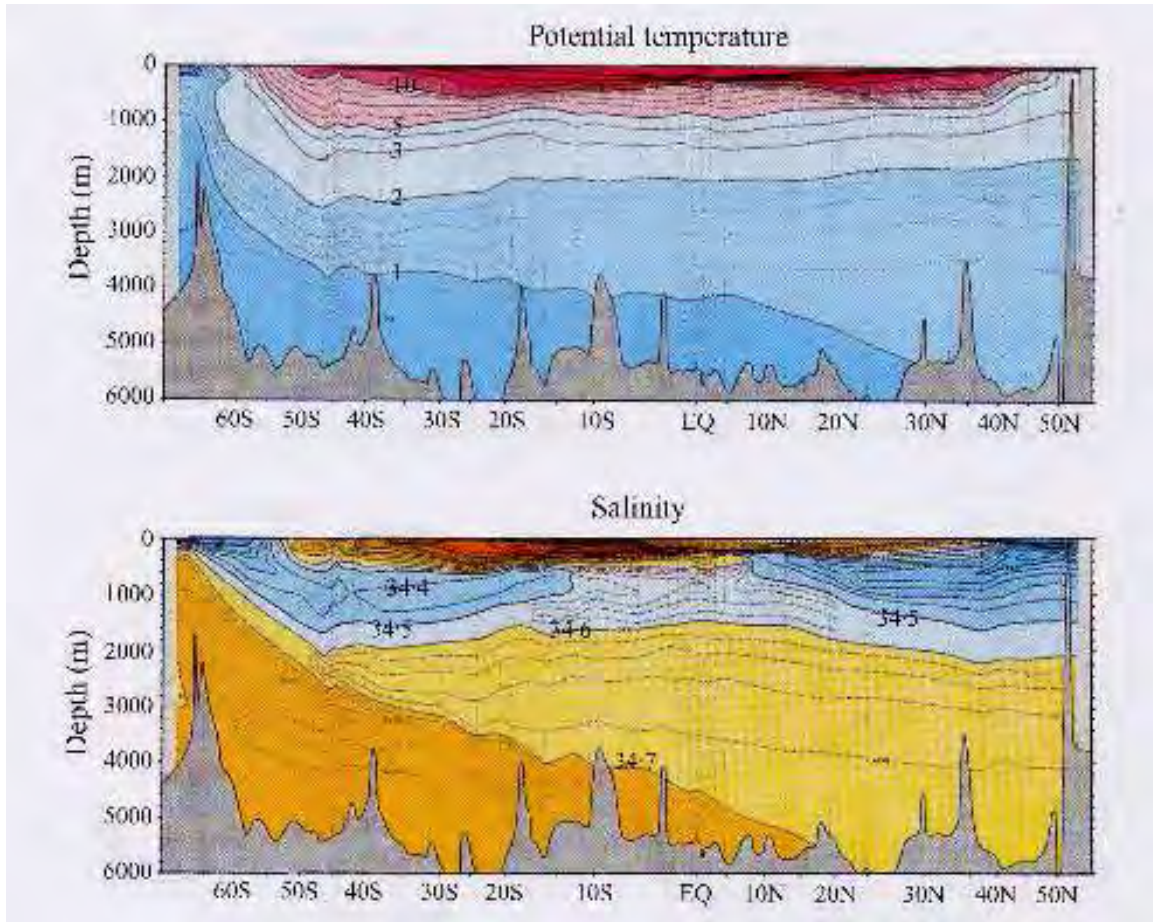


Figure 1: GEOSECS section of the potential temperature and salinity in a meridional cross section of the Pacific. The observed stratification, combined with estimates of the circulation, can be used to infer how much mixing must occur.

- Inference - Balancing budgets for confined regions such as deep basins, using measurements of the inflow together with simple models to infer parameters;
- Measurements - for instance, injecting dye or measuring small scale variations in temperature and velocity;
- Process Studies - modeling and quantifying the physical processes such as internal wave formation and breaking which give rise to the mixing.

If models aim to represent realistic physics and therefore be able to predict future behaviour, it is important that they take account of the processes rather than just using inferred or measured values. Having the right spatially varying *numbers* is important, but for future prediction we need *formulae* too, so that we take account of any feedbacks in the system.

1.2 Inference

Measurements of temperature and salinity through the depth of the ocean suggest that mixing occurs at all depths, not just in the well mixed surface boundary layer. This results in a stratified interior of the ocean where, in the absence of mixing one would expect an almost uniform core with a much shallower boundary region near the surface where all the mixing occurs. The essential dynamics of turbulent mixing in upwelling regions might be expressed loosely as

$$w \frac{\partial \bar{C}}{\partial z} = \frac{\partial}{\partial z} \left(K_v \frac{\partial \bar{C}}{\partial z} \right), \quad (2)$$

from which, if we can estimate the upwelling rate w , and measure the profile \bar{C} (be it temperature, density or salinity), we can infer an average value of the diffusivity K_v . This has been done (Munk 1966, Munk and Wunsch 1998), and suggests a global average value of $K_v \simeq 10^{-4} \text{ m}^2 \text{ s}^{-1}$, at least below the top 1 km or so of the ocean. This figure must be a reflection of the energy input to the oceans.

More local estimates for K_v can also be made by measuring the flux and water properties of flow into deep basins in which water properties are also known. This gives an estimate for the rate of mixing which must occur within the basin and allows an average K_v to be calculated. Typically this also gives values on the order of 10^{-5} to $10^{-4} \text{ m}^2 \text{ s}^{-1}$. It is not clear from such integrated estimates, however, *where* the mixing occurs - is it in the ocean interior or does it all occur near the boundaries?

1.3 Measurements

Direct observations of mixing come from dye tracer experiments (Ledwell et al. 1998), in which inert dye was released at a specific depth in the ocean and the vertical spread of this dye over time was measured. This suggests that the diffusivity in the open ocean should be on the order of $K_v \sim 10^{-5} \text{ m}^2 \text{ s}^{-1}$, rather less than the inferred global average, though at shallower depths than those to which the global estimates apply, so that there is not a contradiction.

The kinetic energy dissipated can be calculated from measuring the small scale (<cm) variations in the velocities;

$$\epsilon = \nu |\nabla \mathbf{u}|^2. \quad (3)$$

In the stratified ocean, a fraction of this dissipation, often called the mixing efficiency Γ , is used to increase the potential energy of the fluid by raising the centre of mass of the fluid. Γ is generally taken to be around 0.2. The energy transferred to potential energy is also related to the diffusivity by

$$\Gamma\epsilon = K_v N^2, \quad N^2 = -\frac{g}{\rho} \frac{\partial \rho}{\partial z}, \quad (4)$$

so that if detailed measurements of the velocity on the small scale can be made, a value for the diffusivity K_v can be inferred. However it is unclear if Γ should be a universal constant or should depend upon conditions; in particular it may depend on the stratification N compared with the typical frequency of the waves and also, perhaps, on a parameter describing the strength of wave breaking.

The diffusivities inferred from measuring ϵ are around $10^{-5} \text{ m}^2 \text{ s}^{-1}$ in the open ocean, but show large spatial variations and are more like $10^{-4} \text{ m}^2 \text{ s}^{-1}$ in the deep ocean over regions of rough bottom topography. This suggests that the inferred globally averaged abyssal diapycnal mixing of several $10^{-4} \text{ m}^2 \text{ s}^{-1}$ may be concentrated in localised regions near boundaries or above rough topographies.

1.4 Biological considerations

There was a suggestion from Bill Dewar (Dewar et al. 2006) that a significant proportion of the mixing in the ocean might be caused by the movement of marine animals, particularly zooplankton, which are known to travel up and down depths of more than 1 km every day to feed and escape predators. Large numbers of zooplankton, which tend to move in clouds several hundreds of metres wide, could generate turbulence on a large scale and cause mixing well below the surface mixed layer. It is estimated that around 63 TW is generated by marine animals, and that around 1% of this could be transferred to mechanical energy. Given that overall the mixing of the oceans is thought to require around 2 – 3 TW, this may not be an insignificant contribution, though it has so far been viewed with a degree of scepticism.

1.5 Internal waves

The greater mixing occurring over rough regions of the bed may indicate that it is the result of internal wave interactions; the waves are generated by flow over the bed and propagate up into the ocean core, where non-linear interactions result in overturning and turbulent mixing.

A method developed by Henyey et al. (1986) attempts to use internal wave theory to infer dissipation rates from measurements of shear and strain over length scales of 10 m. This requires some assumptions on scale separation, but seems to produce the expected results reasonably well in the open ocean. It allows ϵ and hence K_v to be estimated over much larger regions, since the velocity need only be measured at 10 m rather than cm intervals. The results are typically consistent with $K_v \approx 10^{-5} \text{ m}^2 \text{ s}^{-1}$.

The general impression built up from these various inferences and measurements is that energy for mixing is radiated to the ocean interior through internal waves generated both at the surface (by the wind) and at the bottom (by tidal currents over rough topography;

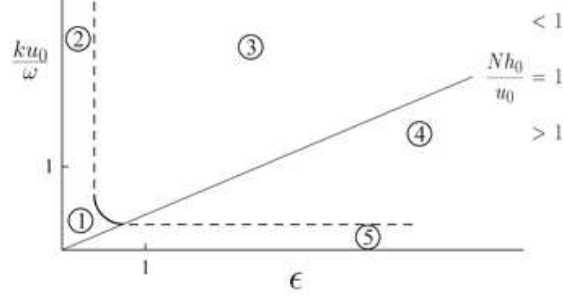


Figure 2: Parameter space for deep-ocean internal tide generation. ϵ is the ratio of bottom slopes to the internal ray slope, and ku_0/ω is the ratio of the tidal excursion to the roughness scale. See Garrett and Kunze, 2007.

internal tides have been visualised as surface waves near Hawaii using satellite measurements (Egbert and Ray, 2000)). Wave-wave interactions set up horizontal currents which lead to overturning and eventually turbulent mixing. Questions arise as to how much of this energy comes from the different sources, and how to parameterise its transfer to the mixing regions.

The energy flux from the wind into near-inertial waves which can propagate below the surface mixed layer varies seasonally and is very dependent on latitude - most of the energy input from this source is at mid latitudes, particularly over the Southern Ocean (Alford 2001, Park et al. 2005). The input is estimated through a combination of satellite and drifter data and numerical models, and it is found that on the order of 1 TW is transmitted from the wind to near-inertial motion.

The energy derived from the barotropic tides has been estimated using satellite data by Egbert and Ray (2001). The tidal elevations can be measured and, through comparison with model simulations, the amount of damping at the ocean floor can be inferred; this is found to occur not just in shallow water off the coasts but also in the open ocean, particularly in regions of rough bottom topography. The estimated power that can be transferred to internal wave generation by this method is around 1 TW globally. It should be noted that Egbert and Ray's method is not accurate in shallow water and may give misleading results in shallow coastal areas.

1.6 Generating internal waves from the tide

The generation of internal tides (that is, internal waves of tidal frequency ω) from barotropic flow over rough topography $h(\mathbf{x})$ depends on several key parameters (Garrett and Kunze, 2007); the ratio ku_0/ω of the length of tidal excursions to horizontal roughness scale, the ratio ϵ of bottom slopes to the internal ray slope, and the ratio of topographic height to the ocean depth. For a deep ocean, there are several different regimes (figure 2) depending on whether none, both or one of ϵ and ku_0/ω are small. The easiest to analyse is when ϵ is small, allowing a linear theory (Bell, 1975). When ku_0/ω is small there are singularities at critical bed slopes, where $\epsilon = 1$ (Balmforth et al. 2002).

Many areas of the ocean such as the Mid-Atlantic Ridge have subcritical slopes and ϵ is small, so the linear theory can be applied. Most of the energy flux into internal tides is in low modes (St. Laurent and Garrett, 2002). With the linear theory, the vertical energy flux

from the bottom can be relatively easily calculated, and for a sinusoidal bed $h = h_0 \sin kx$ is

$$F = \frac{1}{4} \frac{\rho}{\omega} [(N^2 - \omega^2)(\omega^2 - f^2)]^{1/2} k u_0^2 h_0^2. \quad (5)$$

An interesting question one could pose is, supposing this energy flux is dissipated (by internal wave breaking, causing mixing) over some vertical length scale (which could itself depend on the energy flux and N), what happens as the stratification (measured by N^2) changes? Since the energy dissipation causes mixing which alters the stratification, it is of interest whether increasing N^2 causes more mixing which reduces the stratification and would therefore be stable, or whether it causes less mixing, which increases the stratification and causes an unstable positive feedback.

1.7 Transfer of energy and internal wave breaking

The bulk of the energy of internal waves at the M_2 tidal frequency (semi-diurnal) is in low modes which can propagate over long distances from their generation site before breaking up by interactions with each other or with fixed boundaries. It is estimated that up to 1 TW of the 3.5 TW tidal power dissipated by the ocean could be ‘available’ by this method to cause mixing in the ocean interior. The fact that the energy can be generated in one place and transported to another place before dissipating makes it difficult to parameterise. It will be important to understand where and how this mixing occurs; whether the energy flux in low modes breaks down into shorter waves through wave-wave interactions or by bouncing off rough topography, or whether it is transported to continental slopes where it is possible that the energy could be ‘wasted’ in mixing the water in a bottom boundary layer which is already well mixed.

The take home message is that mixing parameterisations (for large scale numerical models this essentially means eddy diffusivities) cannot be based only on purely local conditions. There is much spatial variability which can be measured and included, but understanding and parameterising the *processes* which give rise to this is important if models hope to be truly predictive.

2 Energy from the sea?

2.1 Introduction

The oceans play a major role in global energy issues, both in terms of their current uses and role in the carbon cycle, and in their potential to be exploited as an alternative energy source in the future. Currently large amounts of oil and gas are recovered offshore, ocean water is used for cooling nuclear power plants (and the oceans were once used for nuclear waste disposal), and there is an increasing number of offshore wind farms being developed.

Wave power has been suggested as a means of extracting energy from the surface of the ocean but, apart from small local projects, will probably not be able to produce energy on a significant scale. Tidal power has somewhat more potential for larger scale projects and probably offers more opportunities as a significant ‘new’ power source. Assessing how much energy can be extracted from such projects and where they should be placed is therefore of great interest. There may also be the possibility to use larger scale circulatory currents,

such as the Gulf Stream, to supply useful amounts of power, and the feasibility of such schemes poses a large range of fluid dynamical questions.

2.2 A global perspective

The tides caused by the motion of the moon and sun are estimated to supply 3.7 TW (3.7×10^{12} W) of power which must be dissipated by transfer to kinetic, potential and internal energy of the earth, atmosphere and oceans (Munk and Wunsch, 1998). The vast majority of this energy transfer, approximately 3.5 TW, occurs in the oceans, initially as surface tides, with energy eventually being dissipated through bottom friction and internal wave formation leading to localised turbulence and mixing. As a comparison, human power usage is around 15 TW, a large power station generates around 1 GW, the total wind dissipation in the atmosphere is around 1000 TW, and insolation provides nearly 10^5 TW (advances in solar power are presumably therefore the most promising way forward).

An important first question when we consider large scale tidal power generation is how much of this we can feasibly tap into without altering the tidal dynamics so much that we start to reduce the available power. A simple model to assess this question was suggested; that is to treat the Earth's oceans as a simple harmonic oscillator with a natural resonant frequency ω_0 which is near to the frequency of tidal forcing ω :

$$\ddot{x} + (\lambda_0 + \lambda_1)\omega_0\dot{x} + \omega_0^2x = F(t) \equiv \cos \omega t. \quad (6)$$

The natural dissipation of the oceans is written as the damping force $\lambda_0\omega_0\dot{x}$, and the term $\lambda_1\omega_0\dot{x}$ is the extra dissipation which we might impose through tidal power generation. $F(t)$ is the tidal forcing at frequency ω . The average power dissipated from the system is

$$P_{total} = \overline{(\lambda_0 + \lambda_1)\omega_0\dot{x}^2} = \overline{F\dot{x}}, \quad (7)$$

where the overbars denote the average over a cycle. The power generated for human use is a fraction of this;

$$P_1 = \frac{\lambda_1}{\lambda_0 + \lambda_1} P_{total}, \quad (8)$$

and the question is how does this vary as we vary λ_1 ?

The solution of (6) is

$$x(t) = \frac{(\omega_0^2 - \omega^2) \cos \omega t + (\lambda_0 + \lambda_1)\omega\omega_0 \sin \omega t}{(\omega_0^2 - \omega^2)^2 + (\lambda_0 + \lambda_1)^2\omega^2\omega_0^2}, \quad (9)$$

giving

$$P_{total} = \frac{1}{2} \frac{(\lambda_0 + \lambda_1)\omega_0\omega^2}{(\omega_0^2 - \omega^2)^2 + (\lambda_0 + \lambda_1)^2\omega^2\omega_0^2}, \quad (10)$$

and

$$P_1 = \frac{1}{2} \frac{\lambda_1\omega_0\omega^2}{(\omega_0^2 - \omega^2)^2 + (\lambda_0 + \lambda_1)^2\omega^2\omega_0^2}. \quad (11)$$

As more artificial dissipation (λ_1) is added at first it produces more power, but eventually the system saturates and then as more friction is added the power output decreases. P_1 has

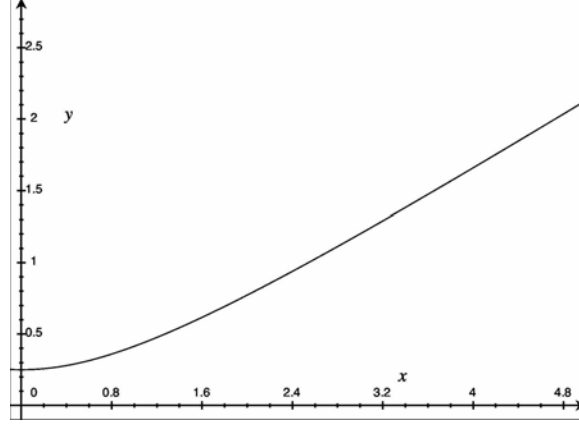


Figure 3: Fraction of current tidal dissipation which could be used for human energy sources as a function of the distance from resonance x of the tides, from (14). Dominant tidal modes are likely to be close to resonance, so the relevant section of this graph is where x is small and the ratio is close to $1/4$.

a maximum when $(\omega_0^2 - \omega^2)^2 + (\lambda_0 + \lambda_1)^2 \omega^2 \omega_0^2 = 2(\lambda_0 + \lambda_1)\lambda_1 \omega_0^2 \omega^2$, and the power then generated is

$$P_{max} = \frac{1}{4\lambda_0\omega_0} \frac{1}{1 + (1 + x^2)^{1/2}}, \quad (12)$$

where

$$x = \frac{\omega_0^2 - \omega^2}{\lambda_0\omega_0\omega}, \quad (13)$$

is a parameter which describes how far the tidal forcing is from the resonant frequency. Ocean tides are thought to be close to resonance, so that x is a small parameter. This compares with the natural tidal dissipation P_{nat} (that is, (10) with $\lambda_1 = 0$) by a factor

$$\frac{P_{max}}{P_{nat}} = \frac{1 + x^2}{2[1 + (1 + x^2)^{1/2}]}, \quad (14)$$

which is symmetric in x so that we have only considered positive values in figure 3.

Near resonance, x is small and the largest amount we can hope to extract is around $1/4$ of the naturally dissipated energy (3.5 TW). The tides therefore offer a considerable potential source of energy if we can find the engineering capabilities and suitable locations to extract this energy.

2.3 Tidal power engineering

Conventional tidal power projects make use of an existing bay by building a barrage across the bay entrance which traps the water in the bay at high tide. Once the level of the water outside the bay has fallen the water is released through turbines, the energy coming from the drop in potential energy of the water. The bay is thus filled and emptied during every tidal cycle. If the tidal range is a depth $2a$ and the bay has a surface area A_s , then the

maximum potential energy release as the water passes through the turbines is

$$\frac{1}{2}\rho_w g A_s (2a)^2, \quad (15)$$

and since this can only occur once per tidal cycle (maybe twice if the process is repeated on the flood tide, but we ignore this possibility) the average power produced is

$$\frac{\omega}{\pi}\rho_w g A_s a^2, \quad (16)$$

where $\omega \approx 2\pi/12.4 \text{ h}^{-1}$ is the frequency of the dominant M_2 tide. For a reasonably large basin of area $A_s = 100 \text{ km}^2$ and a tidal range of $2a = 4 \text{ m}$ this gives about 240 MW. The La Rance facility in France is an example of this type of project in operation and several other possible sites have been mooted. Naturally the obvious place to build such projects is in areas where the tidal range is very large such as the Bay of Fundy in Eastern Canada and the Severn estuary in the UK. Unfortunately this property means that these locations are also considered important ecological sites and there is much debate as to whether the benefits of developing large schemes can justify the loss of unique habitats and ecosystems.

Another possibility, alongside or instead of using bays and estuaries, is to place turbines in channels where there are strong tidal currents and to use the existing currents to drive the turbines. Many such places have been suggested and the ‘rights’ to develop tidal schemes in some such locations are being claimed (in Puget Sound for instance). Of course, these schemes also have potential ecological problems as well as posing problems for shipping and fishing.

2.4 Extracting power from a tidal channel

As our ability to build and position large submarine turbines increases it is important to be able to assess where the best place to position the turbines is, and how many should be used in order to make the most efficient use of the available energy. Just as on the global scale, if too many turbines are used in one location the tidal currents will be reduced and the power output will decrease. Maximising the power output from a channel means optimising the number and position of turbines. If channels converge or diverge it is important to consider what effect building a turbine in one location will have on the flow in other channels, and similarly if currents are going to be utilised in conjunction with damming bays or estuaries then the interaction between the different schemes needs to be carefully considered.

The power available from any given channel is often misleadingly estimated by measuring the kinetic energy flux through a cross-section. If flow at velocity u occurs along a channel of area A between two large basins then the kinetic energy flux through the cross-section is

$$\frac{1}{2}\rho_w A u^3. \quad (17)$$

This will be a misleading value to quote as the power which could be extracted from a particular location since it is highly dependent on where in the channel it is measured and there is no reason to assume that, if a turbine is placed at the location with the greatest flux, the same power will be generated by the turbine. Moreover, if more than one turbine is

placed in the same channel one clearly cannot hope to extract the same energy flux twice or more! Interpreting maps of potential turbine sites must therefore be done with due caution.

The problem of maximising the power output from flow down a channel can be considered most simply by looking at the effect of a ‘fence’ of turbines constructed across the channel (whether this would be feasible for shipping etc. is questionable, but calculations suggest that a complete fence may be the most efficient option for power generation).

For flow along a channel of cross section $A(x)$ at velocity $u(x, t)$, a one dimensional force balance for the flow is

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + g \frac{\partial \zeta}{\partial x} = -F, \quad (18)$$

in which ζ is the surface elevation and F is the friction which may be written in terms of a drag coefficient C_D as

$$F = \frac{C_D u^2}{h} + F_{turb}. \quad (19)$$

F_{turb} is the additional friction introduced by the turbine, and the quantity of interest is the power dissipated by the turbine,

$$P = \rho_w \overline{Q F_{turb}}, \quad (20)$$

where the overbar denotes a space-time average. Q here is the volume flux Au , and if we assume that the channel is short compared to the wavelength of the tide (likely to be hundreds of kilometres) then volume conservation requires that this volume flux is independent of distance along the channel.

We imagine the flow in the channel is driven by the changing elevation of the basins at either end and we write $\Delta\zeta(t)$ as this head difference (note that this could vary between positive and negative over the tidal cycle). We can think of prescribing this head difference, e.g. $\Delta\zeta = a \cos \omega t$, as the tidal forcing in this model. With $u(x, t) = Q(t)/A(x)$, we integrate (18) over the length of the channel L to get

$$c \frac{dQ}{dt} = g \Delta\zeta - \alpha Q |Q| - F_{turb}, \quad (21)$$

where

$$c = \int_0^L \frac{1}{A} dx, \quad \alpha = \frac{1}{2} \frac{1}{A_L^2} + \int_0^L \frac{C_D}{h A^2} dx, \quad (22)$$

A_L is the cross-sectional area at the outlet, where the flow separates, and we assume that the flow at the inlet is simple sink flow so that there is no additional term from there.

The question of interest is what happens to the generated power P as we vary the friction provided by the fence of turbines F_{turb} . In general this requires knowing how this turbine drag depends on the flow rate (e.g. a linear or quadratic dependence), and (21) must then be solved to find the maximum power according to (20). A simple quasi-steady limit can be considered; if the acceleration is unimportant then the turbine-free channel has

$$Q(t) = \left(\frac{g \Delta\zeta(t)}{\alpha} \right)^{1/2}. \quad (23)$$

When a turbine is introduced, Q will differ from this, but how it differs will depend on the form of F_{turb} . In the quasi-steady state the power can be written

$$P = \rho_w \overline{Q(g \Delta\zeta - \alpha Q |Q|)}, \quad (24)$$

and we can therefore find the maximum possible power available by finding the Q which makes this largest. This happens when

$$Q(t) = \left(\frac{g\Delta\zeta(t)}{3\alpha} \right)^{1/2}, \quad (25)$$

and the power is then

$$P = \frac{2}{3^{1/2}} \rho_w g \Delta\zeta(t) \left(\frac{g\Delta\zeta(t)}{\alpha} \right)^{1/2} \approx 0.21 \rho_w g a Q_{max}, \quad (26)$$

if $\Delta\zeta = a \cos \omega t$ and $Q = Q_{max} = (ga/\alpha)^{1/2}$. In this case Q is reduced to $1/3^{1/2} \approx 0.58$ of its natural value and $2/3$ of the head drop $\Delta\zeta$ is associated with the operation of the turbines.

Including acceleration and using different forms for the turbine drag gives the general result that

$$P = \gamma \rho_w g a Q_{max} \quad (27)$$

with the numerical factor γ somewhere between 0.19 and 0.24 (Garrett and Cummins, 2005).

This analysis suggests that if a fence of turbines is built, the location along the channel is unimportant, provided there is only one such fence. This is in contrast to the simplistic power estimate (17); we can compare the above result with that estimate in the case when the drag coefficient C_D is small so that from (26),

$$P = \frac{2}{3^{1/2}} \rho_w \alpha \overline{\left(\frac{g\Delta\zeta(t)}{\alpha} \right)^{3/2}} \approx \frac{2}{3^{1/2}} \frac{1}{2} \rho_w A_L \overline{u_L^3}, \quad (28)$$

where u_L is the exit velocity in the natural state (23). The predicted maximum power is a fraction $2/3^{1/2} \approx 0.38$ of that predicted naively from the kinetic energy flux at the channel outlet. In general, it is likely that (17) will give an overestimate of the available power.

Detailed numerical calculations have been made to check some of these assumptions and to assess the potential in more complex situations such as the Johnstone Strait in British Columbia. Here the flow splits between different tidal passages and introducing a turbine in one will cause more of the flow to be diverted into other passages, thus reducing the maximum power available to the turbine. This emphasises the importance of assessing the impact of tidal schemes over large areas to avoid overstating the potential of these schemes.

2.5 Extracting power from a tidal bay

A very similar analysis can be carried out for flow into and out of a bay, when turbines can be located across the mouth of the bay and increase the friction there. In this case the volume of water in the bay must be conserved so that the height of the water in the bay $\zeta(t)$ will evolve separately, with a phase lag, from the forcing height $\zeta_0(t)$ outside the bay. A corresponding model might be

$$c \frac{dQ}{dt} = g(\zeta_0 - \zeta) - \alpha Q |Q| - F_{turb}, \quad A_s \frac{d\zeta}{dt} = Q, \quad (29)$$

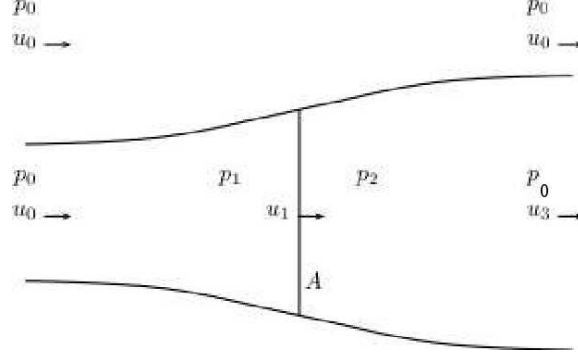


Figure 4: Steamtube passing through a turbine of cross-section A , across which the flow is u_1 and there is a pressure drop $p_1 - p_2$, so that the drag on the turbine is $F = (p_1 - p_2)A$. The background flow is u_0 .

in which we attempt to maximise the power

$$P = \overline{\rho_w Q F_{turb}}, \quad (30)$$

by choosing the optimum amount of turbine friction.

With some assumptions about the form of the turbine drag, the maximum power is again found to be

$$P = \gamma \rho_w g a Q_{max}, \quad (31)$$

where $\gamma \approx 0.24$ is a factor which might vary between 0.19 and 0.26, a is the tidal amplitude, and Q_{max} is the maximum flow in the undisturbed state. This gives a power 0.76 times the reference value (16) for a conventional basin scheme in a basin with no natural friction in the entrance. This maximum however occurs when the reduction of the tidal range in the bay is 74% (there is also a change in the phase shift between bay and outside). A scheme such as this therefore has substantial benefits for ecosystems since the natural semidiurnal filling and emptying of the bay is maintained with relatively little alteration.

For both tidal channels and bays, we can conclude from these simple analyses that the maximum power obtainable with a fence of turbines is given approximately by (31) with $\gamma = 0.22$ in terms of the natural tidal amplitude a and maximum flow rate Q_{max} .

2.6 Fences vs. individual turbines

Complete barrages of turbines are probably unfeasible for a number of reasons, so it is of interest what power is lost by using isolated turbines or an array of turbines, when energy is dissipated in the mixing of the wake from each turbine.

The power on an isolated turbine in an infinitely wide channel can be estimated by expressing the force on the turbine in two ways - one using the change in momentum flux across the turbine, the other using Bernoulli's equation for the streamtube which passes through the turbine, as shown in figure 4. This gives

$$p_0 + \frac{1}{2} \rho_w u_0^2 = p_1 + \frac{1}{2} \rho_w u_1^2, \quad p_2 + \frac{1}{2} \rho_w u_1^2 = p_0 + \frac{1}{2} \rho_w u_3^2, \quad (32)$$

whence

$$p_1 - p_2 = \frac{1}{2}\rho_w(u_0^2 - u_3^2), \quad (33)$$

and the force on the turbine of area A is

$$F = \frac{1}{2}\rho_w A(u_0^2 - u_3^2), \quad (34)$$

whilst the change in momentum flux gives

$$F = Au_1(u_0 - u_3). \quad (35)$$

Hence $u_1 = (u_0 + u_3)/2$ and the power generated by the turbine can be written as a multiple of the kinetic energy flux through the undisturbed turbine cross-section:

$$P = Fu_1 = \frac{1}{2}(1+r)(1-r^2) \times \frac{1}{2}\rho_w Au_0^3, \quad r = \frac{u_3}{u_0}. \quad (36)$$

This has a maximum when $r = 1/3$, so $u_1 = 2/3 u_0$, and the power is then $16/27$ of the undisturbed kinetic energy flux, the ‘Lanchester-Betz’ limit. A fence of turbines giving the same head loss would generate power Fu_0 , so the proportion of power missed by the isolated turbine as compared to a fence is $1 - u_1/u_0 = 1/3$. This missing power is lost to dissipation as the slow-moving fluid in the turbine wake merges with the faster free stream.

A similar analysis can be applied to an isolated turbine in a finite width channel and the energy loss compared to a fence increases towards a limit of $2/3$.

2.7 Other forms of oceanic hydroelectricity

An interesting, although unlikely, suggestion was made concerning the damming of the Red Sea, allowing some of the trapped water to evaporate and then using inflow across the resulting head drop at a rate that balances the evaporation rate of about 1.7 m/yr. The insolation over the Red Sea is around 10^5 GW, so this power would seem to be very well worth harnessing. After 10 years of evaporation however, a head of $H = 17$ m could be exploited to produce power $\rho_w g H Q \approx 4.1$ GW. This is much less than the power causing evaporation, a result of the large latent heat of evaporation - it takes as much energy to evaporate water as to raise it 250 km vertically, so that dropping it a mere 17 m does not seem to be a very effective way of harnessing the power!

In conclusion, tidal power may be a useful, though not hugely abundant, energy source. There are many issues to consider; engineering feasibility, ecological changes, shipping, and careful positioning of schemes are vitally important. The possibility of using tidal currents around headlands or using giant turbines to extract the energy tied up in the global oceanic circulation have been suggested, though not explored fully. It is not clear what impact turbines will have on such ‘unenclosed’ currents and how effective they would be at harnessing the available power. It is also not clear what impact the increased friction would have upon such currents. It may be the case that less environmental impact is caused by making use of tidal currents in other ways, for instance as an efficient means of cooling nuclear power stations.

References

- [1] Alford, M H 2001 Internal swell generation: The spatial distribution of energy flux from the wind to mixed layer near-inertial motions. *J. Phys. Oceanogr.* **31**:2359–2368
- [2] Balmforth, N J, Ierley, G R and Young, W R 2002 Tidal conversion by subcritical topography. *J. Phys. Oceanogr.* **32**:2900–2914
- [3] Bell, T H 1975 Topographically generated internal waves in the open ocean. *J. Geophys. Res.* **80**:320–327
- [4] Bell, T H 1975 Lee waves in stratified flows with simple harmonic time dependence. *J. Fluid Mech.* **67**:705–722
- [5] Dewar, W K, Bingham R J, Iverson R L, Nowacek D P, St. Laurent L C and Wiebe P H 2006 Does the marine biosphere mix the ocean? *J. Mar. Res.* **64**:541–561
- [6] Egbert, G D and Ray, R D 2000 Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data. *Nature* **405**:775–778
- [7] Egbert, G D and Ray, R D 2001 Estimates of M_2 tidal dissipation from TOPEX/Poseidon altimeter data. *J. Geophys. Res.* **106**:22475–22502
- [8] Garrett, C and Cummins, P 2005 The power potential of tidal currents in channels. *Proc. R. Soc. A* **461** 2563–2572 doi:10.1098/rspa.2005.1494
- [9] Garrett, C and Kunze, E 2007 Internal Tide Generation in the Deep Ocean. *Annu. Rev. Fluid Mech.* **39**:57–87
- [10] Henyey, F S, Wright, J and Flatté, S M 1986 Energy and action flow through the internal wave field - an eikonal approach. *J. Geophys. Res.* **91**:8487–8495
- [11] Ledwell, J R, Watson, A J and Law, C S 1998 Mixing of a tracer in the pycnocline. *J. Geophys. Res.* **103**:21499–21529
- [12] Munk, W.H. 1966 Abyssal recipes *Deep-Sea Res.* **13**: 707–730
- [13] Munk, W and Wunsch, C. 1998 Abyssal recipes II: energetics of tidal and wind mixing. *Deep-Sea Res.* **45** 1976–2009
- [14] Park, J J, Kim, K and King, B A 2005 Global statistics of inertial motions. *Geophys. Res. Lett.* **32**: L14612, doi:10.1029/2005GL023258
- [15] St. Laurent, L and Garrett, C 2002 The role of internal tides in mixing the deep ocean. *J. Phys. Oceanogr.* **32**:2882–2899

Herbert Huppert: Gravity Currents & Solidification

A. Shravat and S. Barbot

January 12, 2009

This lecture discussed a number of topical problems in geological fluid mechanics. Two particular examples are covered in details in these notes: the dynamics of the grounding line of an ice sheet & shelf, and the fluid mechanics of carbon sequestration. Some further mathematical and physical details are relegated to the appendices.

1 Ice Sheet Grounding Line

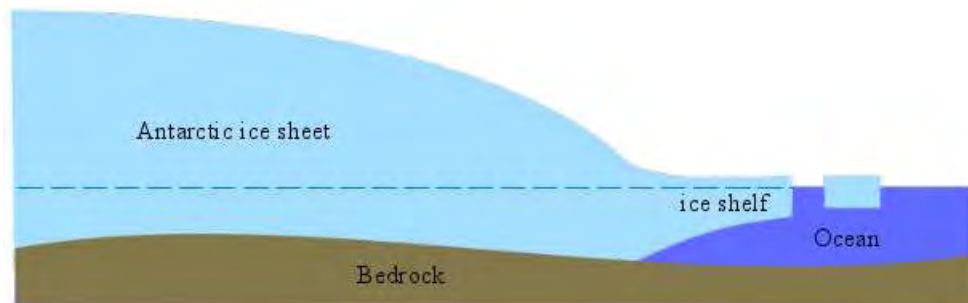


Figure 1: Ice Sheets and Ice Shelves.

In the present day, the last of Earth's large ice sheets can be found in Greenland and Antarctica. These ice sheets are continental in size and are characterized by complex dynamics that may be driven by climate forcing or the spacial and temporal variations at the ice bed or internal boundaries. Ice shelves are enormous beds of floating ice as can be seen in Figure 1 that gives a schematic view of ice sheets and shelves. Figure 2 shows photos of floating ice sheets and icebergs in the sea.

The ice sheet/ice shelf transition zone plays an important role in controlling marine ice sheet dynamics, as it determines the rate at which ice flows out of the grounded part of the ice sheet. Any change in ice thickness will of course lead to migration of the grounding line. There have been several experiments on the dynamics of the grounding line. One such experiment can be seen in Figure 3, where wax is poured over water to create a grounding line.

Ice Sheets and Icebergs



Figure 2: Floating ice sheets and Icebergs.

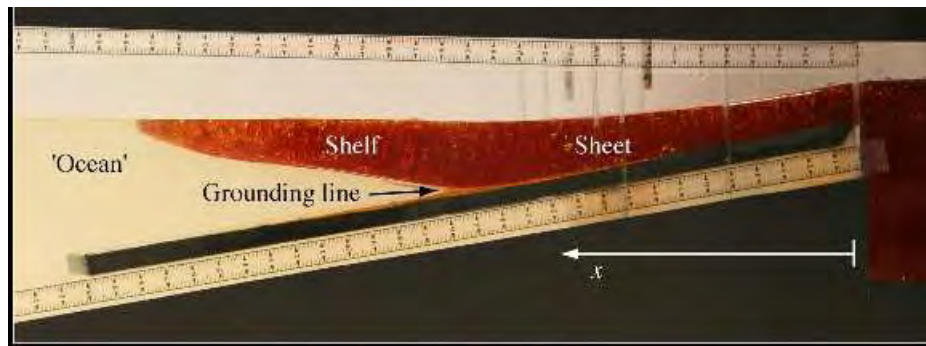


Figure 3: Laboratory experiment with wax showing the formation of the shelf.

1.1 Dynamics of Grounded Ice Sheet

The Grounded Ice Sheet can be modelled as a shear flow (see Figure 4). The dynamical balance is between the pressure gradient and the shear stress.

$$-\rho g \frac{\partial h}{\partial x} \approx \mu \frac{\partial^2 u}{\partial y^2}. \quad (1)$$

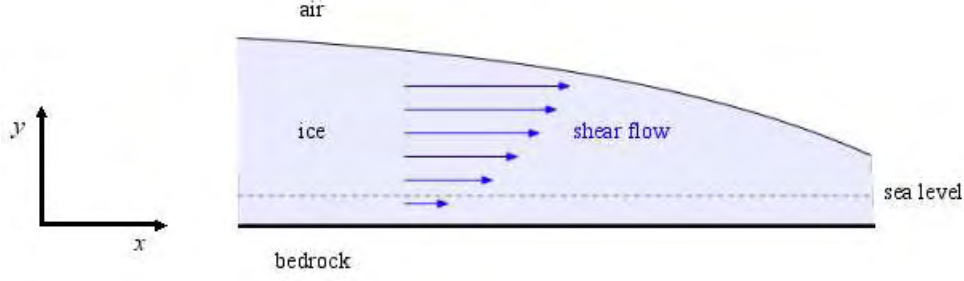


Figure 4: Schematic of the Grounded Ice Sheet.

The horizontal length scale, obtained from dimensional analysis (see section 1 in appendix A) is

$$L \sim \left(\frac{\rho g q^3}{\mu} \right)^{1/5} t^{4/5}. \quad (2)$$

where q is the precipitation rate. The horizontal length is plotted against t in Figure 5.

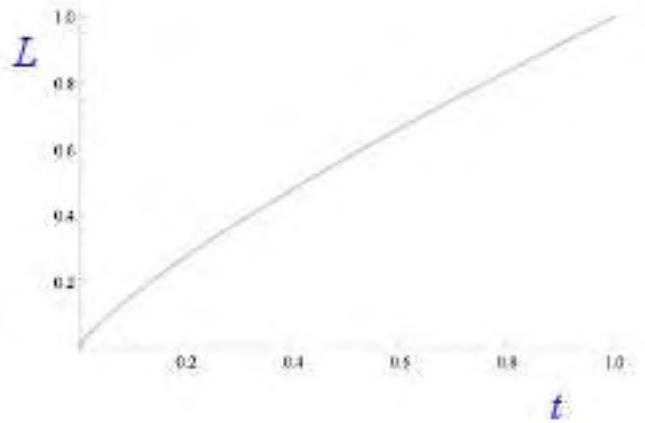


Figure 5: Plot of horizontal length scale, L against t for the grounded ice sheet.

1.2 Dynamics of Freely Floating Shelf

In the preceding section, the sheet has been modelled as fixed on a bedrock. The dynamics of a freely floating shelf shall be considered now. The leading order dynamical balance is satisfied by the pressure gradient balancing the inertia.

$$-\rho g' \frac{\partial h}{\partial x} \approx \frac{4\mu}{h} \frac{\partial}{\partial x} \left(h \frac{\partial u}{\partial x} \right), \quad (3)$$

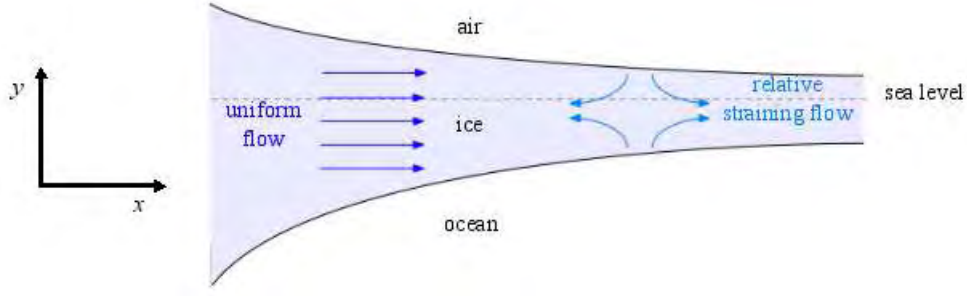


Figure 6: Schematic showing the freely floating shelf.

where $g' = \frac{\Delta\rho}{\rho}g$. The horizontal scale is given by

$$L \sim \frac{\rho g' q}{\mu} t^2. \quad (4)$$

The plot of horizontal scale, L against t is shown in Figure 7.

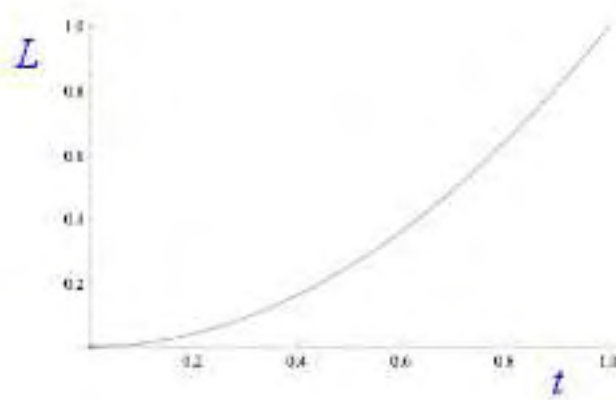


Figure 7: Graph of L against t for freely floating shelf.

1.3 Modelling Ice Sheet-Shelf Dynamics

The dynamics of both ice sheet and shelf together is considered in this section. A sketch of this with various parameters is given in Figure 8. The flux in the sheet is given by

$$q = \frac{gH^3}{3\nu} \left(-\frac{\partial h}{\partial x} \right). \quad (5)$$

Here $H = h + b$. The continuity (mass conservation) requires that

$$\frac{\partial h}{\partial t} = -\frac{\partial q}{\partial x}, \quad (6)$$

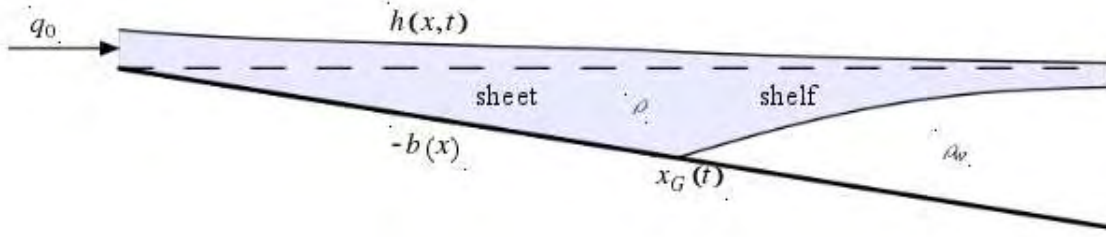


Figure 8: Schematic of the ice sheet and shelf with flux q_0 due to precipitation. The parameter $b(x)$ defines the bottom topography.

This gives rise to the nonlinear diffusion equation,

$$\frac{\partial h}{\partial t} = \frac{\partial}{\partial x} \left[\frac{g}{3\nu} (h + b)^3 \frac{\partial h}{\partial x} \right]. \quad (7)$$

The prescribed volume flux (due to precipitation) at $x = 0$ is given by

$$q_0 = -\frac{g}{3\nu} h^3 \frac{\partial h}{\partial x}, \quad (8)$$

The floatation condition at $x = x_G(t)$ is defined as

$$\rho H = \rho_w b, \quad (9)$$

The normal stress balance at $x = x_G(t)$ is

$$\left(\frac{\Delta\rho}{\rho} \frac{db}{dx} - \frac{\partial h}{\partial x} \right) x_G = \frac{gH^2}{8\nu} \left[4 \left(\frac{\partial h}{\partial x} \right)^2 - \frac{\Delta\rho}{\rho} \right]. \quad (10)$$

Theory gives the following results:

$$x_G(t) \sim t^{1/2}, \quad (11)$$

at early times. For late times,

$$x_G(t) \rightarrow \text{constant}. \quad (12)$$

The various grounding line positions are plotted against time in Figure 9. Theory predicts the position of a steady grounding line to be

$$x_G = \frac{1}{\alpha} \frac{\rho}{\rho_w} \left(\frac{6\nu q_0}{g} \right)^{1/3} \left(\frac{g}{g'} \right)^{1/6}.$$

Figure 10 provides a comparison between the theoretical and experimental results. Figure 11 compares some of the theoretical steady profiles with actual ice sheets from Antarctica.

To summarise, ice sheets and shelves flow as viscous gravity currents. In grounded ice sheets, shear stresses balance the hydrostatic pressure gradient. The result is a decelerating

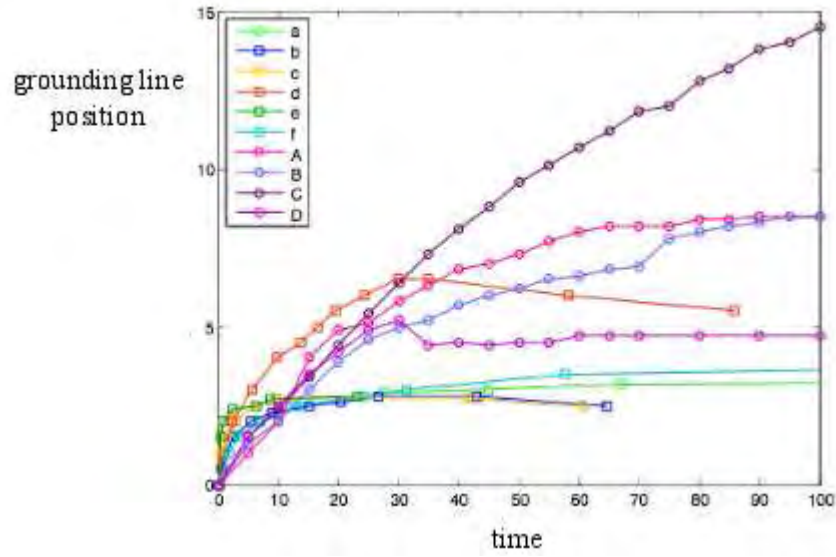


Figure 9: Plot showing the grounding line position as a function of time in the case of ice sheet-shelf dynamics.

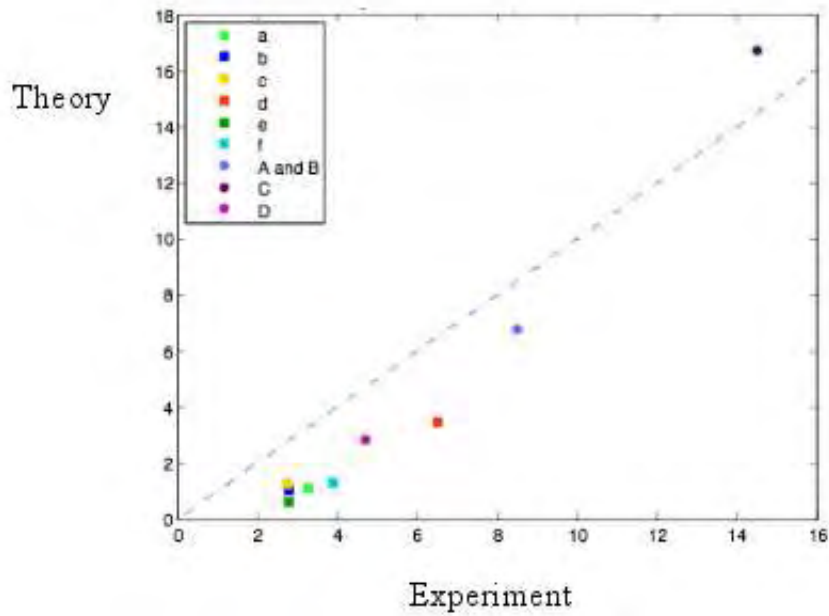


Figure 10: A comparison between theoretical prediction and experimental results for the position of the grounding line.

flow. The floating ice shelf, on the other hand, has extensional stress balancing the hydrostatic pressure gradient, causing the flow to accelerate. This gives rise to a hypothesis that

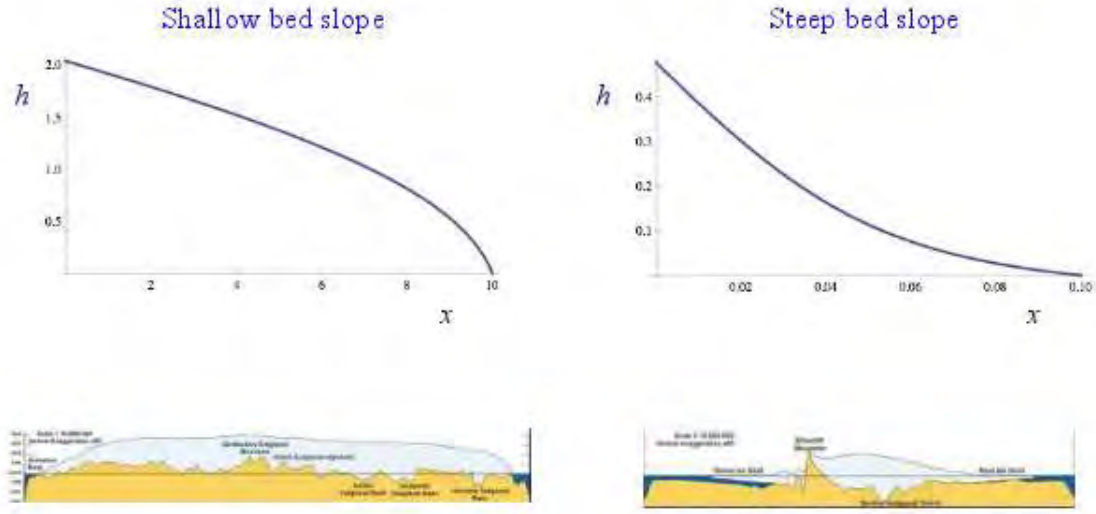


Figure 11: Steady Ice sheets with Grounding Lines.

the dynamic control of grounding line is effected by balance of normal stress. Hence, leading to a closed mathematical model. This hypothesis is being tested using laboratory experiments. One application of this mathematical model is to determine the dynamical stability of ice sheets. As mentioned earlier, grounding line is very sensitive to any change in the climate. To this end, more work is being carried out to understand the basic mechanisms determining its stability and position.

To summarize, ice sheets and shelves can be modelled as viscous gravity currents. In grounded ice sheets, shear stresses balance the hydrostatic pressure gradient. The result is a decelerating flow. The floating ice shelf, on the other hand, has extensional stress balancing the hydrostatic pressure gradient, causing the flow to accelerate. This gives rise to a hypothesis that the dynamic control of grounding line is effected by balance of normal stress. Hence, leading to a closed mathematical model. This hypothesis is being tested using laboratory experiments. One application of this mathematical model is to determine the dynamical stability of ice sheets. As mentioned earlier, grounding line is very sensitive to any change in the climate. To this end, work is being carried out to understand the basic mechanisms determining its stability and position.

2 Fluid Mechanics of Carbon Dioxide Sequestration

2.1 Introduction and Motivation

The Keeling curve (Fig. 12) is a graph showing the variation in concentration of atmospheric carbon dioxide since 1958. It is based on continuous measurements taken at the Mauna

Loa Observatory in Hawaii under the supervision of Charles D. Keeling from the Scripps Institution of Oceanography. His measurements show evidence of CO_2 undergoing a regular seasonal cycle, reflecting the seasonal growth and decay of land plants in the northern hemisphere, as well as a regular long-term rise driven by the burning of fossil fuels. Before the industrial era, atmospheric CO_2 concentration was between 275 and 280 ppmv for several thousand years. Carbon dioxide has risen continuously since then, and the average value when Dr. Keeling started his measurements in 1958 was near 315 ppmv. By the year 2000 it has risen to about 367 ppmv. Keeling's measurements showed the first significant evidence of rapidly increasing carbon dioxide levels in the atmosphere.

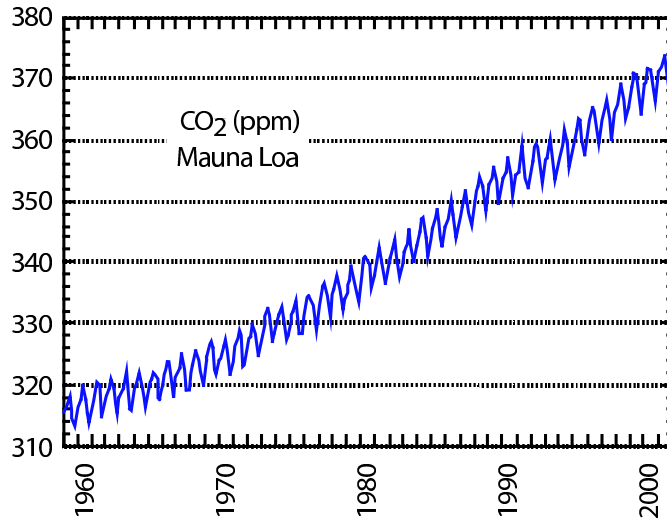


Figure 12: Keeling Curve. Atmospheric carbon dioxide concentration (ppm) above Mauna Loa observatory in Hawaii.

Fig. 13 compares the world and the United States population and energy usage in 1906 and 2006. At the start of the industrial era, the world population was about one billion souls. It took 120 years to double this number. Today, at the rate about 200,000 new people per day, it takes 13 years for the world population to grow by a billion. Today's large human population and rate of growth is not without effect on the environment. Between 1906 and 2006, the U.S. as well as the world population grew nearly four fold. The world energy usage increased almost 13-fold. In the U.S. alone, energy usage in 2006 is 9 times its value in 1906.

Fig. 14 shows the source contributions of world energy consumption. Fossil fuels (coal, oil and natural gas) still account for over 85% of the primary energy consumed in the world. Fossil fuels or mineral fuels are fossil source fuels, that is, hydrocarbons found within the top layer of the Earth's crust. The burning of fossil fuels produces around 21.3 billion tons of carbon dioxide per year, but it is estimated that natural processes can only absorb about half of that amount, so there is a net increase of 10.65 billion tons of atmospheric carbon dioxide per year (one ton of atmospheric carbon is equivalent to $44/12$ or 3.7 tons of carbon dioxide). Carbon dioxide is one of the greenhouse gases that enhances radiative forcing and contributes to global warming, causing the average surface temperature of the Earth to rise

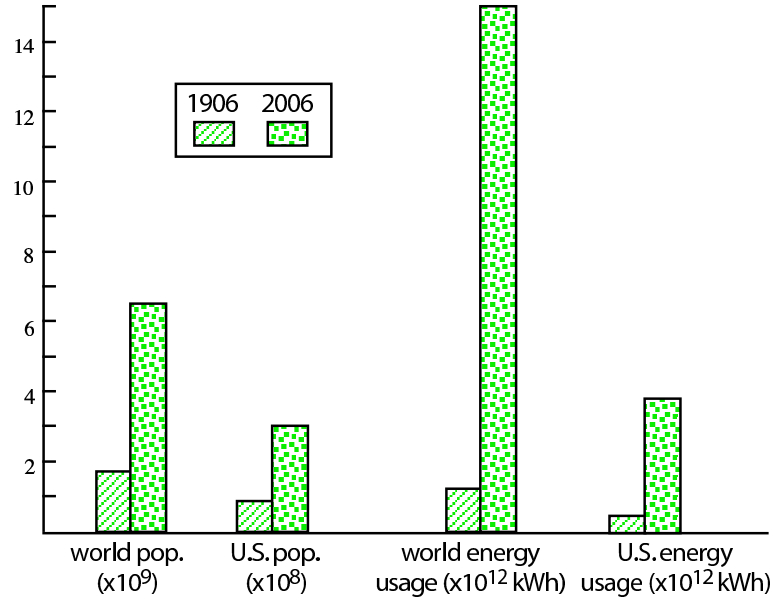


Figure 13: A hundred years evolution of the United States and the world's population and energy usage. Between 1906 and 2006, the U.S. as well as the world population grew nearly four fold. The world energy usage increased almost 13-fold. In the U.S. alone, energy usage in 2006 is 9 times its value in 1906.

in response.

In the last 400,000 years temperature was highly correlated with CO_2 concentration in the atmosphere. Inferred mean temperature anomalies (difference with present time) exhibit a roughly 100,000 year cycle with amplitude varying from -6°C to $+2^\circ\text{C}$. Effect of strong increase of CO_2 in the atmosphere is associated with climate change. Long records of carbon dioxide concentration in the atmosphere indicate that CO_2 content never exceeded 300 ppm in the last 400,000 years. The amplitude of the sudden increase of CO_2 concentration in the last hundred years, associated with burning of fossil fuel, is unprecedented in recent geological history. The effects, mechanisms and time scales of climate change are subject to active research. Limiting the effects of human activities on the environment involves a migration to renewable energies and the reject of fossil fuel as a primary source of energy. One way to reverse the Keeling curve, however, to return to geologically normal atmospheric carbon dioxide concentration, is to the re-injection of CO_2 in underground reservoirs. The so called carbon dioxide sequestration provides means to remove CO_2 from the atmosphere and restore the status quo of pre-industrial era. The problem of carbon dioxide sequestration calls for further investigation of the behavior of gravity (buoyancy) driven flows in porous media.

2.2 Source in Porous Medium

Flow in porous media occurs in many natural and industrial situations. Included in these are the seepage of rainwater through permeable ground into an aquifer, the forced flow of

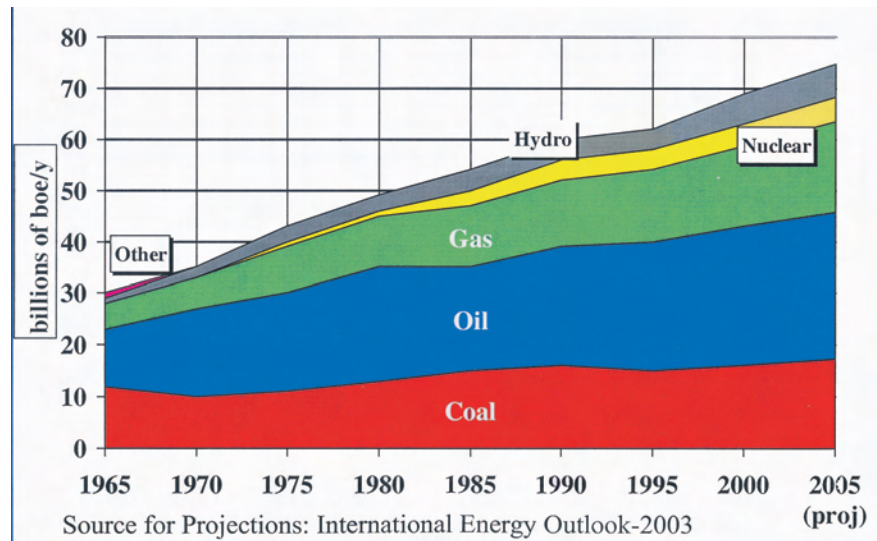


Figure 14: Distribution of energy source over the last five decades. Fossil fuels still account for over 85% of the primary energy consumed in the world.

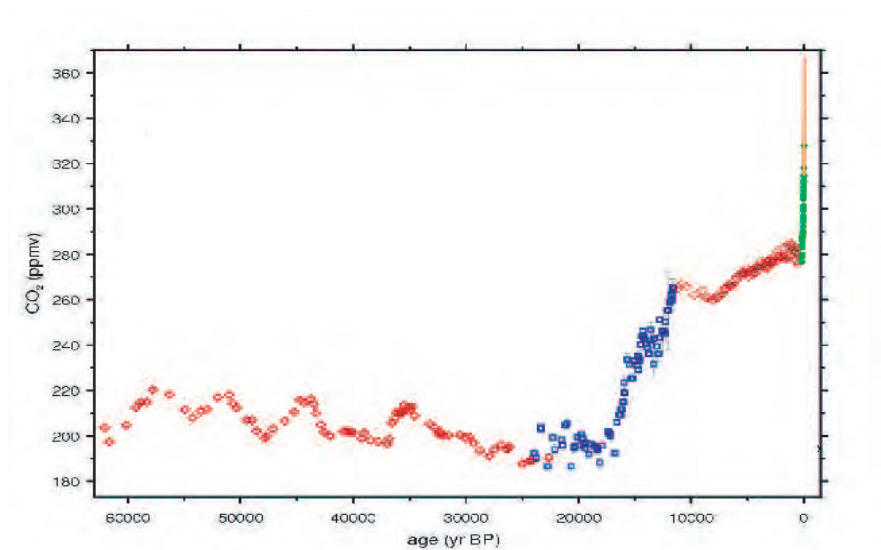


Figure 15: Carbon dioxide levels over the last 60,000 years in volume ppm. Longer records show that CO₂ concentration did not exceed 300 ppm in the last 400,000 years. Observed increase in CO₂ content in the last hundred years is unprecedented in recent geological history.

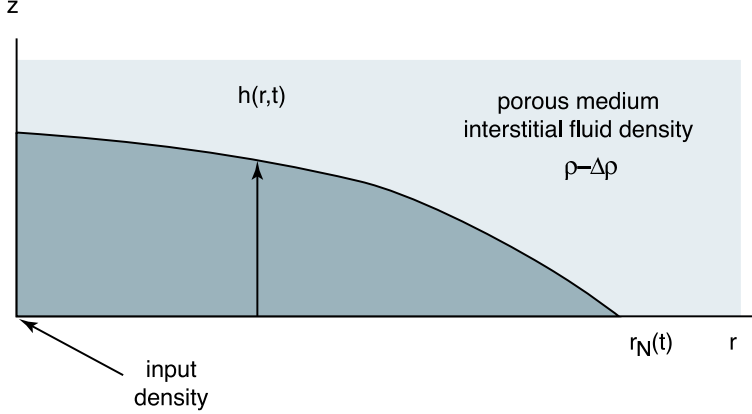


Figure 16: Schematic of gravity fluid in a porous medium.

oil from sandstone reservoirs (Lake 1989; Gerritsen & Durlofsky 2005), and the dispersion of polluted liquids through gravel pits. In some situations the fixed, solid matrix through which the flow passes reacts with the interstitial fluid and the structure and porosity of the matrix change with position and time (Hallworth, Huppert & Woods 2005). Examples of such reactions include the gradual formation of dolomite (Phillips 1991) and the convective flow due to solidification through a mushy layer, which is a region of reactive solid matrix bathed in interstitial fluid (Worster 2000; Aussillous et al. 2006).

A gravity current occurs whenever fluid intrudes primarily horizontally into fluid of different density. The fundamentals of flows beneath a relatively less dense homogeneous fluid layer at either low (Huppert 1982a, b, 2000) or high Reynolds number (Benjamin 1968; Houtt 1972) are well known. In recent years additional phenomena due to the effects of rotation (Ungarish & Huppert 1998), and flows over porous media (Acton, Huppert & Worster 2001; Thomas, Marino & Linden 2004), into stratified ambients (Ungarish & Huppert 2002; Maxworthy et al. 2002) and over variable topography have been investigated.

The spreading of a liquid phase in a porous medium hosting pore fluid of different density, as sketched in Fig. 16, is a kind of gravity current. Consider the gravity current due to horizontal pressure gradient of (unknown) free surface slope (see also Appendix A). The dynamics of the fluid is simplified to the viscous approximation (Darcy's flow)

$$0 = -\nabla p - \rho \mathbf{g} - \mu \mathbf{u}/k \quad (13)$$

where k is the porous medium permeability, μ is the dynamic viscosity of input fluid and \mathbf{u} is the input fluid velocity. In the horizontal direction, this becomes

$$\frac{\partial p}{\partial r} = \rho g' \frac{\partial h}{\partial r} \quad (14)$$

where

$$g' = g \frac{\Delta \rho}{\rho} \quad (15)$$

is the reduced gravity. The equation of local continuity gives

$$\phi \frac{\partial h}{\partial t} + \frac{1}{r} \frac{\partial}{\partial r}(r u h) = 0 \quad (16)$$

where ϕ is the porosity of the host rock. Equating the last two equalities gives

$$\frac{\partial h}{\partial t} - \frac{\gamma}{r} \frac{\partial}{\partial r} \left(r h \frac{\partial h}{\partial r} \right) = 0 \quad (17)$$

and

$$\gamma = \frac{\rho k g'}{\phi \mu} \quad (18)$$

has the dimensions of a velocity. The equation of global continuity gives

$$Q t = 2\pi \int_0^{r_N(t)} r h dr \quad (19)$$

The similarity variable

$$\eta = (\gamma Q / \phi)^{-1/4} r t^{-1/2} \quad (20)$$

leads to an expression for the position of the nose

$$r_N(t) = \eta_N (\gamma Q / \phi)^{1/4} t^{1/2} \quad (21)$$

Defining $y = \eta(r, t) / \eta(r_N, t)$ with $0 < y < 1$ and

$$h(r, t) = \eta_N^2 (Q / \phi \gamma)^{1/2} f(y) \quad (22)$$

one obtains the ordinary differential equation

$$(y f f')' + \frac{1}{2} y^2 f' = 0 \quad (23)$$

with $f(1) = 0$ and

$$\eta_N = \left[2\pi \int_0^1 y f(y) dy \right]^{-1/4} \quad (24)$$

For a linear case, with $Q = 0$, approximation solution of the differential equation gives

$$f(y) \sim \frac{1}{2}(1 - y) \quad (25)$$

and

$$\eta_N \sim (6/\pi)^{1/4} \quad (26)$$

Data from laboratory experiments are well explained by such scaling. Further experiments on gravity currents in porous media may help us predict the fate of liquid phase carbon dioxide injected in porous reservoirs (Fig. 17).

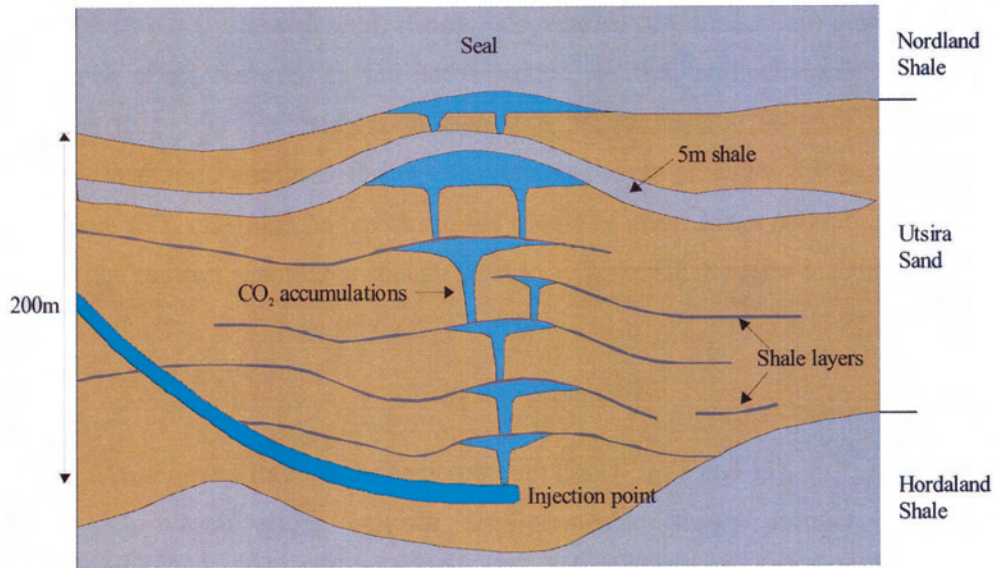


Figure 17: Schematic of carbon dioxide (liquid phase) sequestration and propagation in a porous reservoir

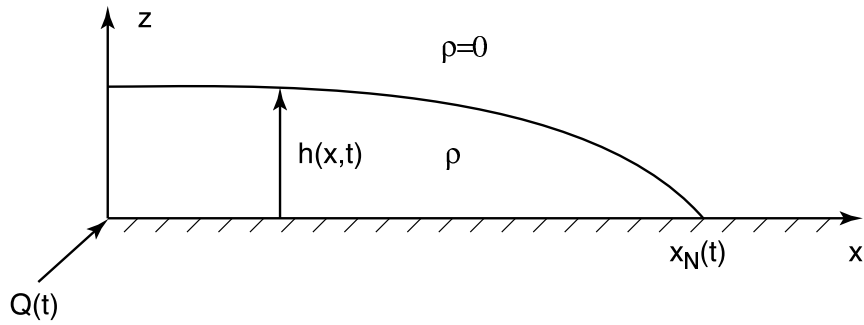


Figure 18: Gravity current model. The flow is driven by pressure gradient due to unknown free surface slope and resisted by viscosity.

Appendices

A Gravity Currents

Gravity currents are generated by the release of some volume of one homogeneous fluid into another of slightly different density. If a viscous fluid is released vertically on a horizontal surface it rapidly takes up a circular plan form as it spreads. This form is observed to be stable to any small disturbances which are initiated on the front due, for example, to irregularities in the horizontal surface or to chance perturbations (Huppert 1980, Huppert 1982a). In this note, we describe how gravity currents spread, and quantify their radius or extent as a function of time.

We consider that the flow is driven by pressure gradient due to unknown free-surface slope $h(x, t)$ and resisted by viscous effects. We first examine a two-dimensional flow propagating in the x -direction only, with a vertical dependence. Model geometry is illustrated in Fig. 18. Assuming the velocity field $\mathbf{u} = u_x(x, z) \mathbf{e}_x$, the Navier-Stokes equation can be written

$$\mu \frac{\partial^2}{\partial z^2} u = -\frac{\partial}{\partial x} p = \rho g \frac{\partial}{\partial x} h \quad (27)$$

where we have ignored inertial effects (Reynolds number $\text{Re} \ll 1$), μ is the dynamic viscosity and p is pressure. Also assumed is a thin layer where $u_{zz} \gg u_{xx}$.

A.1 Dimensional analysis

As shown in Fig. 18, the flow is sustained by the mass flux $\rho Q(t)$. Experimental data indicates the time dependence

$$H L \sim \int_0^t Q(t') dt' \sim q t^\alpha \quad (28)$$

where H , L are typical height and typical length and α is a constant. For $\alpha = 0$, the volume is constant. The case $\alpha = 1$ corresponds to constant flux. Using dimension analysis, one can write orders of magnitude of the total buoyancy force F_g

$$\begin{aligned} F_g &= \iiint_V \frac{\partial p}{\partial x} dx dy dz \\ &= -\rho g \iiint_V \frac{\partial h}{\partial x} dx dy dz \\ &\sim \rho g H^2 W \\ &\sim (\rho g q^2 W / L^2) t^{2\alpha} \end{aligned} \quad (29)$$

and the total viscous force F_ν

$$\begin{aligned} F_\nu &= \mu \iiint_V \frac{\partial^2 u}{\partial z^2} dx dy dz \\ &\sim \mu ULW/H \\ &\sim \mu q^{-1} L^3 W t^{-\alpha-1} \end{aligned} \quad (30)$$

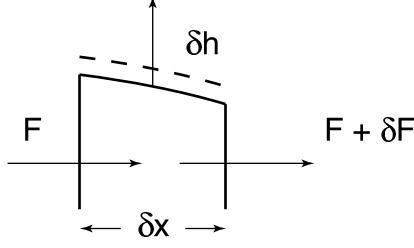


Figure 19: The continuity equation. The change in velocity flux \mathcal{F} is balanced by the change in unknown elevation h .

The buoyancy and viscous forces balance ($F_\nu \sim F_g$) leading to an estimate of the total extension of the gravity flow

$$L = (g q^3 / \nu)^{1/5} t^{(3\alpha+1)/5} \quad (31)$$

A.2 Lubrication theory analysis

We use the approximations of lubrication theory, where we neglect both surface tension and contact line effects (Hocking 1981). Both can be neglected if the Bond number

$$B = \rho g l^2 / T \gg 1 \quad (32)$$

where ρ is the density of the fluid, l is a representative length scale of the current, and T is the surface tension. Using the x -momentum equation, and the thin layer approximation $u_{xx} = 0$, one obtains

$$\frac{\partial^2}{\partial z^2} u = \frac{\rho g}{\mu} \frac{\partial h}{\partial x} \quad (33)$$

We consider the boundary conditions

$$\begin{aligned} u &= 0, \quad \text{at } z = 0 \\ u_z &= 0, \quad \text{at } z = h \end{aligned} \quad (34)$$

The solution velocity is

$$u(x, z, t) = -\frac{g}{2\nu} \frac{\partial h}{\partial x} z(2h - z) \quad (35)$$

where $\nu = \mu/\rho$ is the kinematic viscosity. The velocity flux is

$$\mathcal{F} = \int_0^h u dz = -\frac{1}{3} \frac{g}{\nu} h^3 \frac{\partial h}{\partial x} \quad (36)$$

The velocity flux has the dimensions of $\mathcal{F} \sim L^2 T^{-1}$. Another constraint of the problem is the conservation of mass. The local continuity requires that the difference in velocity flux between to neighboring columns is balanced by the change of height

$$\partial h \partial x + \partial \mathcal{F} \partial t = 0 \quad (37)$$

or

$$\frac{\partial \mathcal{F}}{\partial x} + \frac{\partial h}{\partial t} = 0 \quad (38)$$

Using equation (36), one obtains

$$\frac{\partial h}{\partial t} - \beta \frac{\partial}{\partial x} \left(h^3 \frac{\partial h}{\partial x} \right) = 0 \quad (39)$$

where $\beta = g/3\nu$. Equation (39) corresponds to the nonlinear diffusion equation. The global continuity equation can be written

$$\int_0^{x_N(t)} h(x, t) dx = q t^\alpha \quad (40)$$

where $q t^\alpha$ corresponds to the cumulative (with time) influx of material volume. The method used by H^2 to determine a similarity solution is to write down the governing equation in terms of dimension. Equation (39) can leads to

$$\frac{h}{t} \sim \frac{\beta h^4}{x^2} \quad \text{or} \quad h^3 \sim \frac{x^2}{\beta t} \quad (41)$$

now assuming no influx of mass ($\alpha = 0$), and a conserved area A , one has $h x_N \sim A$. So one gets

$$\frac{x^5}{\beta A^3 t} \sim 1 \quad (42)$$

which suggests the following dimensionless quantity

$$\eta = (\beta A^3)^{1/5} x t^{-1/5} \quad (43)$$

to be suitable for similarity variable. One obtains the estimate

$$x_N \sim (\beta A^3)^{1/5} t^{1/5} \quad (44)$$

Introducing η_N , the value of η at the nose ($x = x_N$) of the current, and using

$$h \sim \frac{A}{x_N} \sim \left(\frac{A^2}{\beta} \right)^{1/5} t^{-1/5} \quad (45)$$

we use the change of variable

$$h(x, t) = \eta_N^{2/3} (A^2/\beta)^{1/5} t^{-1/5} \phi(\eta/\eta_N) \quad (46)$$

and we further define $y \equiv \eta/\eta_N$ with $0 < y < 1$. The partial derivatives become

$$\begin{aligned} \partial_t &= d_t - \frac{1}{5} \frac{y}{t} d_y \\ \partial_x &= \frac{y}{x} d_y \end{aligned} \quad (47)$$

For example, one has

$$h_t = \eta_N^{2/3} (A^2/\beta)^{1/5} \left\{ \frac{1}{5} t^{-4/5} \phi + y t^{-4/5} \phi' \right\} \quad (48)$$

The partial differential equation governing the height of the gravity current becomes the ordinary differential equation

$$(\phi^3 \phi')' + \frac{1}{5} \phi' + \frac{1}{5} \phi = 0 \quad (49)$$

with the boundary conditions $\phi = 0$ and regular at $y = 1$. Integrating twice, we find

$$\phi(y) = \left(\frac{3}{10} \right)^{1/3} (1 - y^2)^{1/3} \quad (50)$$

and $\eta_N = 1.411$. We find that the problem is not solvable analytically for $\alpha \neq 0$. In general however, one has

$$x_N = \eta_N (\beta q^3)^{1/5} t^{(3\alpha+1)/5} \quad (51)$$

Above theory agrees with experiments.

B Solidification

Solidification is the growth of a solid from a liquid phase. It is an important processing route for metals and alloys and has a number of important geological problems associated with it. In most geological problems, it is assumed that the magma has a well-defined melt temperature at which the phase change from liquid to solid occurs. The problem of solidification can be defined by Stefan's condition. The Stefan condition for heat flux \mathbf{q} and velocity of solidification V_n is given by

$$\rho_s L V_n = [\mathbf{n}, \mathbf{q}]. \quad (52)$$

where L is the latent heat per unit mass, ρ_s is the density of the solid and \mathbf{n} is the normal to the surface. Operator $[\mathbf{a}, \mathbf{b}]$ implies the difference between $\mathbf{a} \cdot \mathbf{b}$ on either side of the solid-liquid interface, where a discontinuity in temperature gradient occurs. The location of the phase change boundary or the interface between the solid and liquid is determined as part of the solution. A complication, however, is the changing of this boundary as solidification proceeds. Another difficulty is to keep track of the latent heat of fusion, which is located at the solid-liquid interface as solidification takes place.

B.1 Planar 1-component solidification

Sometimes the solidification occurs in organised planar way. An example of such solidification is that of a horizontal layer of magma that solidifies from its upper surface downwards as a result of being cooled from above. In this case, the overall flow thickness is unimportant in describing the solidification process as long as a molten region is present. The planar 1-component solidification is given by the equation (53), where T is the temperature.

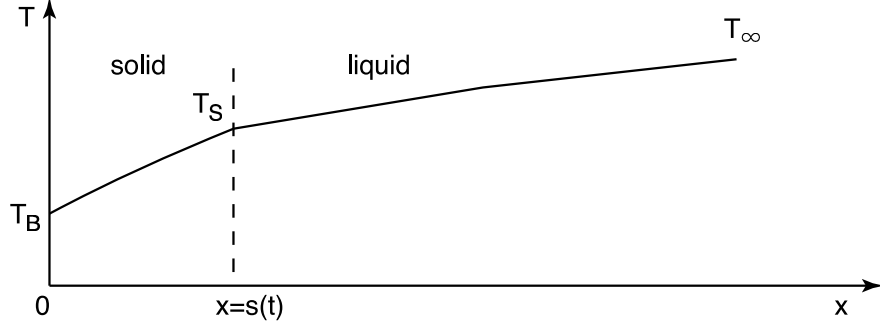


Figure 20: Schematic showing the interface between Solid and Liquid with the boundaries at $x = 0$ and $x = s(t)$. For simplicity material properties are assumed equal.

$$T_t = \kappa T_{xx} \quad (53)$$

where κ is thermal diffusivity. The boundary conditions are:

$$T = T_B \text{ at } x = 0,$$

$$T = T_S \text{ at } x = s(t),$$

$$T \rightarrow \infty \text{ as } x \rightarrow \infty.$$

The Stefan's condition, (52) in this case is given by

$$\rho_s L s' = k(T_x|_{s_-} - T_x|_{s_+}). \quad (54)$$

where k is the thermal conductivity and s' is the first derivative of s with respect to t . This has similarity solution of the form

$$s(t) = 2\lambda(\kappa t)^{1/2}. \quad (55)$$

This can be represented using $\text{erf}(x)$ and λ , which is an eigenvalue of this system defined as a function of S and R . The form of the solution is given by the following expression, where λ_1 is defined as $s_m/2\sqrt{\kappa t}$. s_m gives the solidification interface.

$$\lambda = f[S, R], \quad (56)$$

where

$$S = \frac{L}{c(T_S - T_B)} \quad (57)$$

is the Stefan number relating latent heat to heat capacity and

$$R = \frac{T_\infty - T_S}{T_S - T_B} \quad (58)$$

is the nondimensional thermal forcing. The constant λ_1 is determined by requiring that the latent heat liberated at the solidification boundary be conducted vertically upward, away from the interface.

B.2 Binary alloys

Binary alloys are formed when 2 or more components melt. Sinking of salty water in polar seas or the mixing of iron and impurities in liquid outer core of the Earth are both examples of binary alloys. Figure 21 shows the Phase Diagram for Binary alloys. This representation is often correct in equilibrium but can be very different in many cases. It is worth noting that the freezing temperature (liquidus) is function of composition $T_L(C)$. The solid that freezes has a very different composition (given by solidus) than the liquid. Typically, aqueous solutions have vertical solidus.

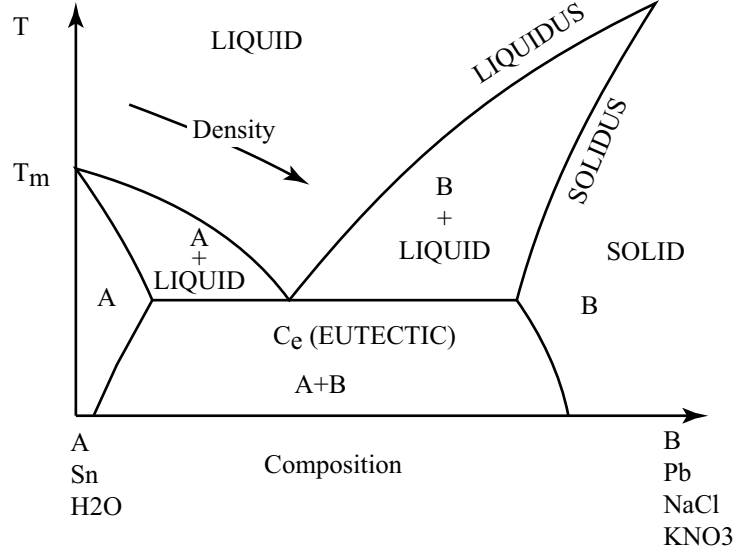


Figure 21: An equilibrium Phase Diagram for Binary Alloys.

B.2.1 Stefan Problem for an alloy

This problem is similar to the Planar 1-component solidification. The equilibrium liquidus can be written as

$$T_i = T_L(C_i). \quad (59)$$

The Stefan condition is

$$\rho_s L s' = -[k T_x]_s^l. \quad (60)$$

The solute conservation is given by

$$(C_i - C_s) s' = -D C_x|_{s_+}. \quad (61)$$

The governing differential can be written in terms of T_t and C_t .

$$T_t = \kappa T_{xx}, \quad (62)$$

$$C_t = D C_{xx}. \quad (63)$$

where D is the diffusion coefficient of C , typically $D \ll \kappa$. Similarity solution can be found of the form

$$s(t) = 2\mu(Dt)^n.$$

The solidification rate is controlled by the concentration diffusion. This is morphologically unstable leading to 'Mushy Layers', which are small scale reactive matrix bathed in interstitial liquid. An example of mushy layer is sea ice. Seeking averaged description of local mean temperature ($T(\mathbf{x}, t)$), mean composition of interstitial liquid ($C(\mathbf{x}, t)$) and solid fraction ($\phi(\mathbf{x}, t)$) leads to the equations for the Mushy layer.

$$\bar{\rho}c(T_t + \mathbf{u} \cdot \nabla T) = \nabla \cdot (\bar{k} \nabla T) + \rho_s L \phi_t, \quad (64)$$

$$(1 - \phi)(C_t + \mathbf{u} \cdot \nabla C) = \nabla \cdot (\bar{D} \nabla C) + (C - C_s) \phi_t. \quad (65)$$

At equilibrium,

$$T = T_L(C).$$

It is safe to say that multicomponent melts solidify into mushy layers, which are often accompanied by strong compositional convection. In many circumstances, it is possible quantitatively to predict structure and rate of growth of mushy layers. An example of solidification is the almost pure iron inner core that has been slowly solidifying for about 1.8×10^9 years (just under half the age of the Earth). During all this time strong compositional convection has maintained the geodynamo protecting us from cosmic radiation.

References

- [1] J. M. ACTON, H. E. HUPPERT & M. G. WORSTER, *Two-dimensional viscous gravity currents flowing over a deep porous medium*, J. Fluid Mech., 440 (2001), pp. 359-380.
- [2] P. AUSSILLOUS, A. J. SEDERMAN, A. GLADDEN, H. E. HUPPERT & M. G. WORSTER, *Magnetic resonance imaging of structure and convection in solidifying mushy layers*, J. Fluid Mech., 552 (2006), pp. 99-125
- [3] T. B. BENJAMIN, *Gravity currents and related phenomena*, J. Fluid Mech., 31 (1968), pp. 209-248.
- [4] M. G. GERRITSEN & L. J. DURLOFSKY, *Modelling fluid flow in oil reservoirs*, Ann. Rev. Fluid Mech., 37 (2005), pp. 211-238.
- [5] M. A. HALLWORTH, H. E. HUPPERT AND A. W. WOODS, *Dissolution-driven convection in a reactive porous medium*, J. Fluid Mech., 535 (2005), pp. 255-285.
- [6] L. M. HOCKING, *Sliding and spreading of thin two-dimensional drops*, Q. J. Mech. Appl. Math., 1 (1981), pp. 37-55.
- [7] D. P. HOULT, *Oil spreading on the ocean*, Ann. Rev. Fluid Mech., 4 (1972), pp. 341-368.
- [8] H. E. HUPPERT, *Geological Fluid Mechanics*. In *Perspectives in Fluid Dynamics: A collective introduction to current research*. Eds. G.K. Batchelor, H.K. Moffatt & M.G. Worster. pp. 447-506 (2000; C. U. P.)

- [9] H. E. HUPPERT, *The propagation of two-dimensional and axisymmetric viscous gravity currents over a rigid horizontal surface*, J. Fluid Mech., 121 (1982a), pp. 43-58
- [10] H. E. HUPPERT, *Flow and instability of a viscous current down a slope*, Nature, 300 (1982b), pp. 427-429.
- [11] H. E. HUPPERT AND J. E. SIMPSON, *The slumping of gravity currents*, J. Fluid Mech., 99 (1980), pp. 785-799.
- [12] L. W. LAKE, *Enhanced oil recovery* (1989; Prentice Hall)
- [13] T. MAXWORTHY, J. LEILICH, J. E. SIMPSON & E. H. MEIBURG, *The propagation of a gravity current in a linearly stratified fluid*, J. Fluid Mech., 453 (2002), pp. 371-394.
- [14] L. P. THOMAS, B. M. MARINO & P. F. LINDEN, *Lock-release inertial gravity currents over a thick porous layer*, J. Fluid Mech., 503 (2004), pp. 291-319.
- [15] M. UNGARISH & H. E. HUPPERT, *The effects of rotation on axisymmetric, particle-driven gravity currents*, J. Fluid Mech., 362 (1998), pp. 17-51.
- [16] M. UNGARISH & H. E. HUPPERT, *On gravity currents propagating at the base of a stratified ambient*, J. Fluid Mech., 458 (2002), pp. 282-301.
- [17] M. G. WORSTER, *Solidification in fluids*. In *Perspectives in Fluid Dynamics: A collective introduction to current research*. Eds. G.K. Batchelor, H.K. Moffatt & M.G. Worster. pp. 393-446 (2000; C. U. P.)

GFD experiments in climate

John Marshall

Notes by: G. Hagstrom & C. Guervilly

January 12, 2009

1 Introduction

This set of notes is about numerical efforts using the MITgcm software package to study the climate on the time scale of thousands of years. Specifically we are interested in three related problems: what sets the pole equator temperature gradient; what determines the total meridional energy transport and its partition between the atmosphere and the oceans; and what determines the extent of the polar ice caps. These are analyzed using numerical calculations on water covered globes which we call aqua planets with a number of barriers to sea flow. The models are simpler than other large scale calculations but retain the essence of the processes that they simulate. The simulations are able to recover various qualitative features of the climate that exist in the real world, including realistic solutions to the aforementioned problems. Furthermore the models are flexible enough to be used to make predictions about certain features of the climate in the distant past or the near future.

2 Meridional Energy Transport From Observations

The temperature of the Earth decreases poleward. The Earth absorbs more energy at the equator than at the poles, and this energy is circulated to the poles where it is radiated away. This processes is known as meridional energy transport. The heat flux through the atmosphere into space has been observed in detail by satellites, and this data can be used to determine the meridional heat flux. Because the flux is so great it must be driven by bulk transport of cold or warm air and water. There is an overall circulating current in the atmosphere and a corresponding one in the ocean with a net transport of warm air polewards and cool air equatorward [3]. Suppose that at a given latitude the mass flux of these currents are Ψ_A and Ψ_O , for the atmosphere and oceans respectively. Then suppose the temperature difference between the currents at each latitude is B_A and B_O . The meridional heat transport at a given latitude would then be:

$$H = H_A + H_O = \Psi_A B_A + \Psi_O B_O. \quad (1)$$

Similarly it is interesting to determine the ratio of the atmospheric and oceanic heat transports:

$$\frac{H_A}{H_O} = \frac{\Psi_A B_A}{\Psi_O B_O}. \quad (2)$$

The partition between atmospheric and oceanic heat transport is known on Earth. These fluxes can be calculated directly by measuring the speed and temperature of currents. Figure 1 is from Trenbeth and Caron [2] and gives the meridional heat transport as a function of latitude. We see that it has a maximum at 35° degrees where it is 6 Petawatts. The total atmospheric heat flux is roughly four times larger than the oceanic flux.

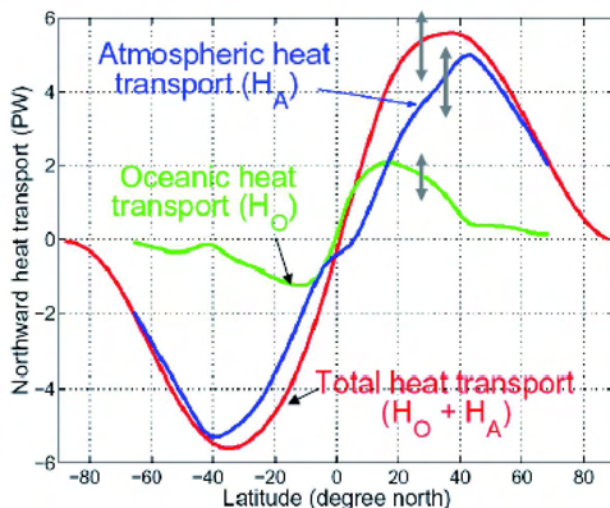


Figure 1: Meridional Heat Transport

The two cases can be analyzed individually, beginning with the oceanic case. The top of the ocean absorbs most of the solar radiation, and as a result is warm. It also has higher salinity due to evaporation. This warm salty layer is driven by the winds. This zone is typically confined to the upper kilometer of the ocean. Underneath it is an abyssal layer which is cold and has lower salinity. The Abyss is vented by convection at the poles, and upwells in the mid and lower latitudes. At its maximum the meridional overturning circulation transports 1.2 Petawatts of heat. The mass flux of this current is on the order of 40 sverdrups.

As is indicated by figure 1 the meridional heat flux is not uniform as a function of latitude. There are also significant variations in the energy contrast and mass flux with latitude. The thermocline is at its most intense in the tropics and the energy contrast is greatest here. In the polar regions the surface temperatures are lower and the ocean is more well mixed.

The meridional circulation of the atmosphere is coupled to that of the ocean. Evaporation of water in the tropics is carried towards the poles. The atmosphere carries moist static energy as a result of its water content. The formula for moist static energy is:

$$B_A = C_A T + gz + Lq. \quad (3)$$

Here C_A is the specific heat of water, g is gravity, z is the altitude, L is the latent heat, and q is the mixing ratio. The moist static energy is the energy that will be released as heat during rainfall. Most of this moisture falls in the midlatitudes. Overall this process accounts for one half of the mass transport of the atmosphere. The differential heating of the atmosphere at the equator and the poles drives this circulation [3].

The atmospheric energy contrast is extremely low in the tropics, so the energy flux must also be low here. On the other hand the energy contrast is quite large in the midlatitudes due to the moist static energy. Figure 2 shows the moisture distribution in the atmosphere as a function of latitude and pressure. The atmosphere is very well mixed in the tropics and this is where most of the convection occurs.

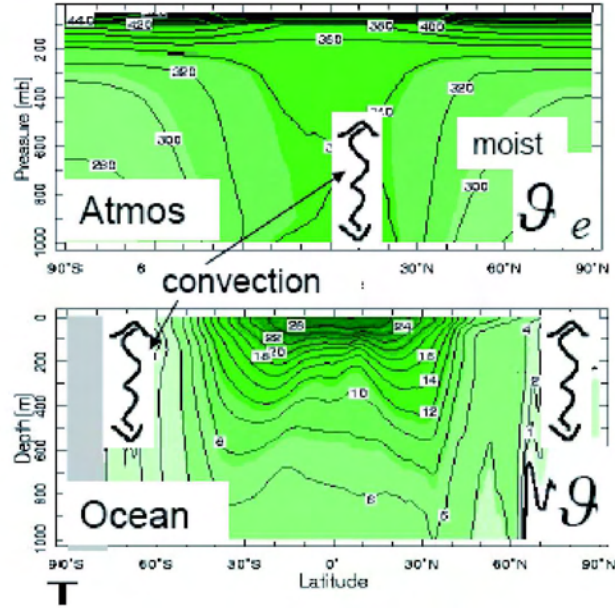


Figure 2: Moisture Distribution in the Atmosphere and Ocean

The qualitative nature of energy transport in the atmosphere and the oceans is quite different. The ocean convects at the poles while the atmosphere convects at the equator. As a result the energy transport is stronger in the ocean in the equatorial regions. The opposite is true everywhere else. On average the atmospheric mass transport is four times greater than that of the ocean, while the energy contrast is the same. These facts are known from observational studies. If we redefine one Sverdrup as 10^9 kilograms per second of any substance then the atmosphere is a one hundred twenty Sverdrup flow while the ocean is only a thirty Sverdrup flow. Figure 3 shows this graphically. The energy contrast and mass transport in the atmosphere are calculated using satellite observations of the mass flux between adjacent moist static energy surfaces. Given these observations it is interesting to wonder what properties of the Earth and its climate cause this partition, and to determine how the partition has been different in the past and how it might change in the future.

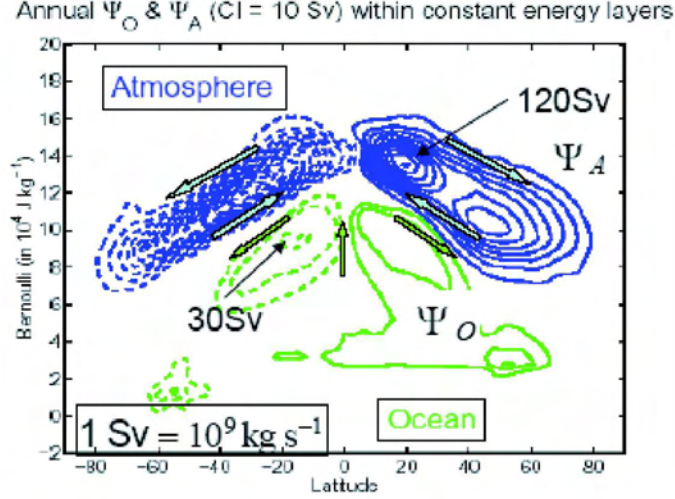


Figure 3: Mass Transport in the Atmosphere and Ocean

3 Modeling Hierarchies and MITgcm

The physics of the oceans and the atmosphere are extremely complex. In many cases the most interesting thing is the large scale behavior. Models of the atmosphere and oceans can range from simple to complex. In the first extreme are box models which divide the ocean and atmosphere into a small number of homogeneous boxes. In the other extreme are large scale general circulation models which discretize the Earth and use numerical methods to approximate the evolution of the atmospheres and oceans. The disadvantage of such complex models is the number of free parameters that can be tuned to cause the model to reproduce observations. Instead of being simulations of the physics these can become simulations of the phenomena. This can obscure the true physical causes for a certain phenomenon and thus make the future predictions of the model unreliable.

Given that scientists have adopted the full range of models for global oceanic and atmospheric circulation it is worthwhile to think about the philosophy of modeling in general. This work shuns the overemphasis on complexity and seeks to use models that correspond to actual physics. The idea is to study a given phenomenon by using a hierarchy of models. One begins with the simplest reasonable model and slowly increases its resolution and features. Here our modeling hierarchy will begin with the aqua planet, which is a planet whose surface is entirely covered with water. To simulate the effect of continents, ridges that block ocean circulation are added to the model, but no additional parameters are introduced. After inserting a ridge we can further increase the analogy by putting a gap in the ridge to represent the Drake Passage. The final step in this study is to consider the double drake, which has two ridges with gaps near the south pole. This hierarchy uses a reasonable model of the physics and should produce interesting results because it captures much of the essence of the geography.

The model that will be used to generate experimental results is MITgcm, which is a general circulation model developed at MIT. It sets itself apart from other models because

of its relative simplicity. It is a realistic physical simulation capable of capturing certain physical process[1].

4 Climate of Aqua Planets

We are interested in qualitative studies of the climate and the ocean circulation using large scale GCM models. The starting point for these studies is the so-called aqua planet. It is a landless rotating planet with a five kilometer deep ocean that is dimensionally modeled after the Earth. Using the data from the simulations we will be able to calculate the partition of energy transport between the atmosphere and the oceans and to compare the qualitative features of this transport. Later when more geography is added to the model it will be possible to create a more nuanced understanding of the global energy circulation. The time scale of each simulation is on the order of a couple of thousand years, and this take one or two weeks of computer time.

Figure 4 shows the sea ice present in the model after it has equilibrated. Ice caps form at each pole and are quite large. The ice reaches a maximum depth of seventeen meters at the actual poles. In the absence of topography the oceans are dominated by circumpolar currents at every latitude. The figure 5 shows the strength of the currents in the atmosphere and ocean at each latitude, depth, pressure, and potential temperature. The currents are quite strong at the equator and then reverse direction at the midlatitudes. The strength drops close to the poles. Combined with figure 6, which shows the humidity and salinity we see that the surface of the ocean in the lower latitudes is warm and salty, whereas the abyss is cooler and fresh. This is very similar to Earth. Similarly in the atmosphere the warm air near the tropics is very moist and the cold air near the poles is quite dry.

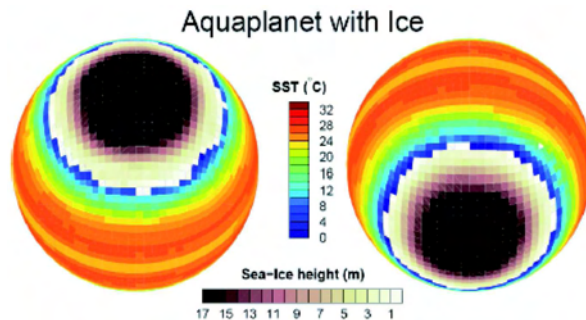


Figure 4: Sea Ice and Water Temperatures in the Aqua Planet

The atmospheric circulation has some qualitative features that exist in the Earth's atmosphere. The beta effect is quite important in driving these. For example figure 7a shows the vertical circulation in both the atmosphere and ocean. There are two clearly defined 'Hadley' cells over the tropics.

One remarkable property of this relatively simple model is that the meridional energy transport is very similar to that of the Earth. Consider the four figures demonstrating the mass transport and energy transport in aqua planet and on Earth, *i.e.* figure 3 and figures 7b, c, and d. The mass transport in the two cases are very close and so is the

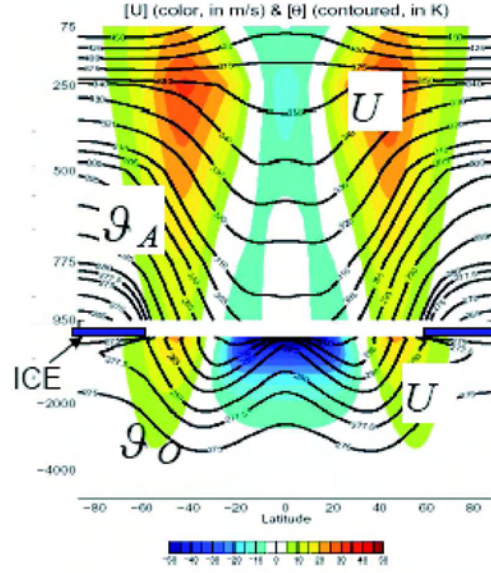


Figure 5: Current Strength and Potential Temperature in the Aqua Planet

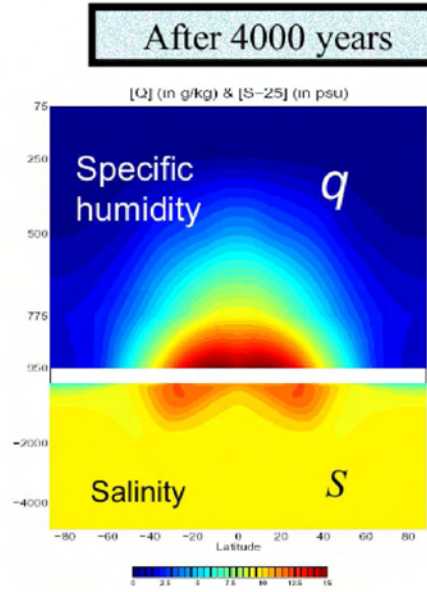


Figure 6: Salinity and Humidity in the Aqua Planet

meridional energy transport. The fact that the shapes of the curves are so similar is astonishing given that the aqua planet has no land. This in many ways represents a victory of the philosophy of modelling hierarchy discussed in the previous section. Although not every feature is captured by this model many of the large scale ones are.

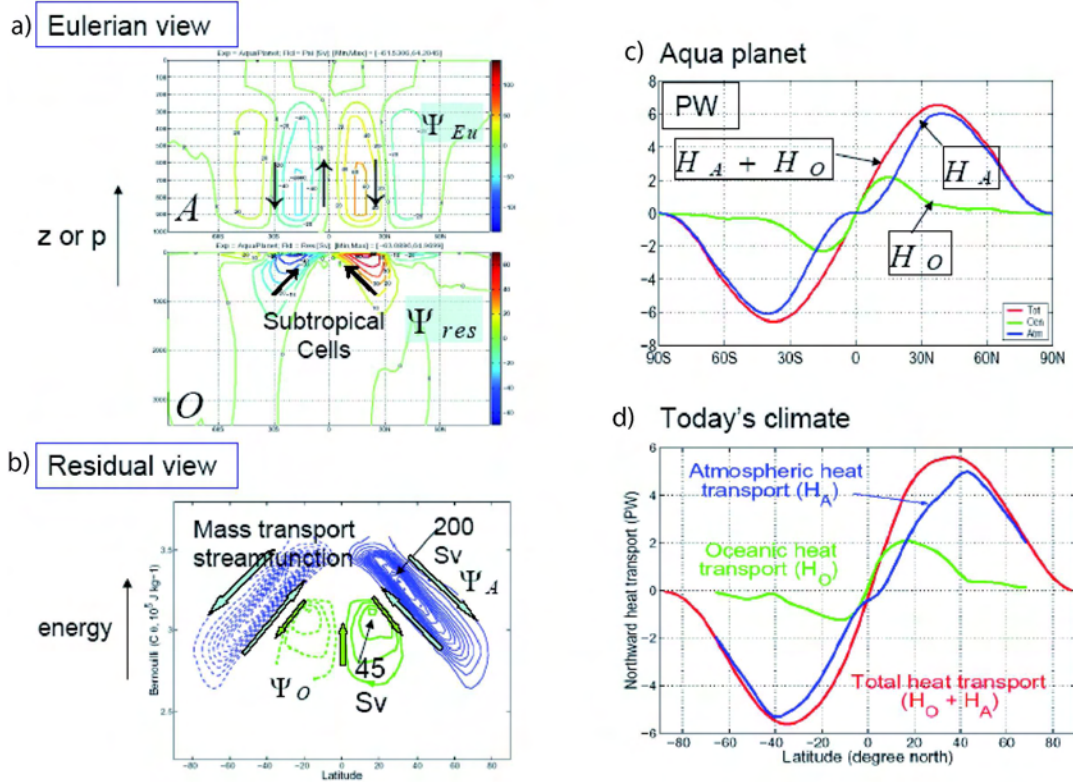


Figure 7: Results from the Aqua Planet: stream function in the ocean and in the atmosphere (a), mass transport (b) and meridional heat transport in Aqua Planet (c) and in Earth (d)

We can use the meridional energy transport data to understand the sea ice caps that are present on the aqua planet. The heat transport drops significantly at the poles. This drop is primarily due to a decline in the mass transport at the poles. In particular the ocean energy transport goes to zero.

5 Role of Geometrical Constraints on Ocean Energy Transport

The coasts are an important constraint for the global circulation in the ocean. The aim of this section is to introduce the impact of the topographic constraints on the flow.

Three variations of the aquaplanet model are discussed (figure 8):

- The ridge model: a ridge going from North pole to South pole blocks the zonal flow in the ocean.
- The drake model: similar to the ridge model but with a passageway representing Drake Passage.
- The double drake model: the ocean is shared into a small and a big basin which can communicate by passageways at the south pole.

In these models, the atmosphere is not blocked by a barrier.

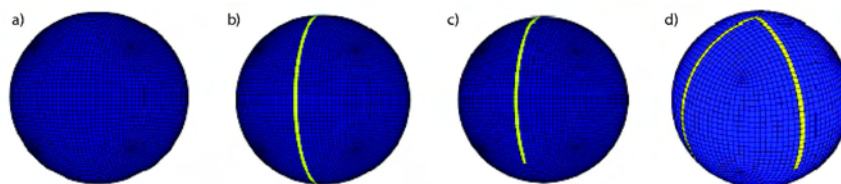


Figure 8: Models with different geometrical constraints: aquaplanet (a), ridge model (b), drake model (c) and double drake model (d).

5.1 Phenomenology

In the ridge model (figure 9a), the oceanic circulation is divided into gyres and a counter current is present along the equator. Moreover, there is no ice cap over the pole.

The passageway in the drake model (figure 9b) allows the presence of a circumpolar current. In this case, unlike the ridge model, an ice cap covers the South pole.

Unlike the aquaplanet, there is no subtropical cell driven by the balance between winds and pressure gradients in the ridge model (figure 10b). A deep convection is present at the two poles. In the drake model (figure 10c), convection is active in the warm hemisphere *i.e.* the northern one where there is no ice. This yields a strong overturning circulation in the northern hemisphere whereas it is tiny in the southern one and then ice can be sustained over the south pole. The salinity is sharply asymmetric in the drake model (figure 11c):

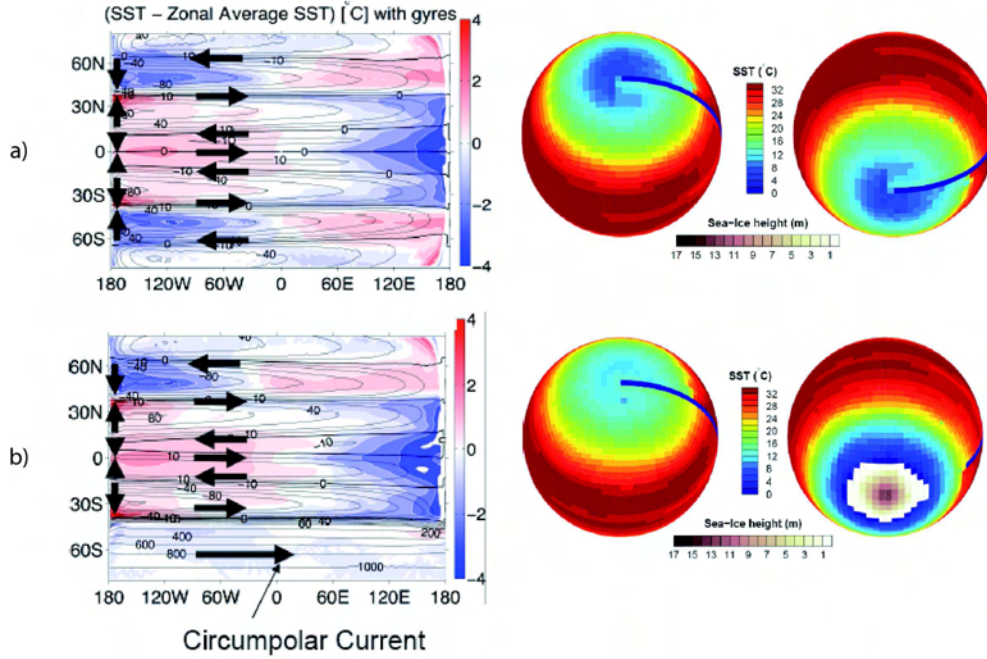


Figure 9: Sea surface temperature computed with the ridge model (a) and the drake model (b).

the salt water is pushed towards the north pole and fresh water towards the south pole. In all models with an oceanic barriers, the atmosphere contains more moisture over the pole than in the aquaplanet model.

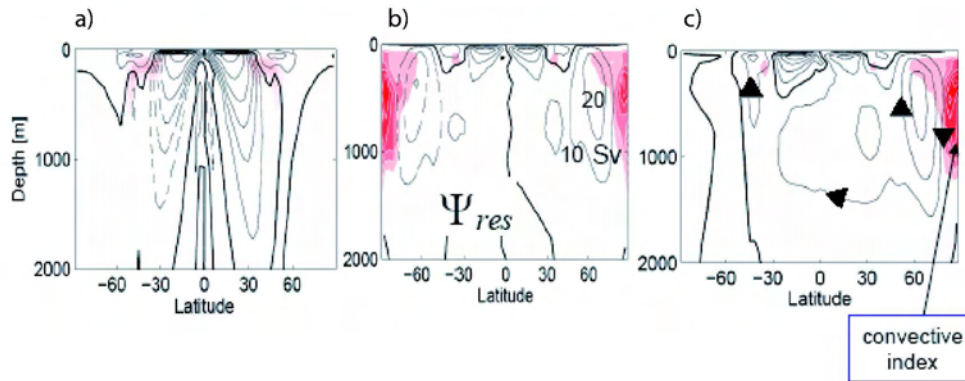


Figure 10: Overturning circulation expressed in sverdrup here redefined as $1\text{Sv} = 10^9 \text{ kg. s}^{-1}$ in the aquaplanet model (a), the ridge model (b) and the drake model (c). The red area represents the intensity of the convection.

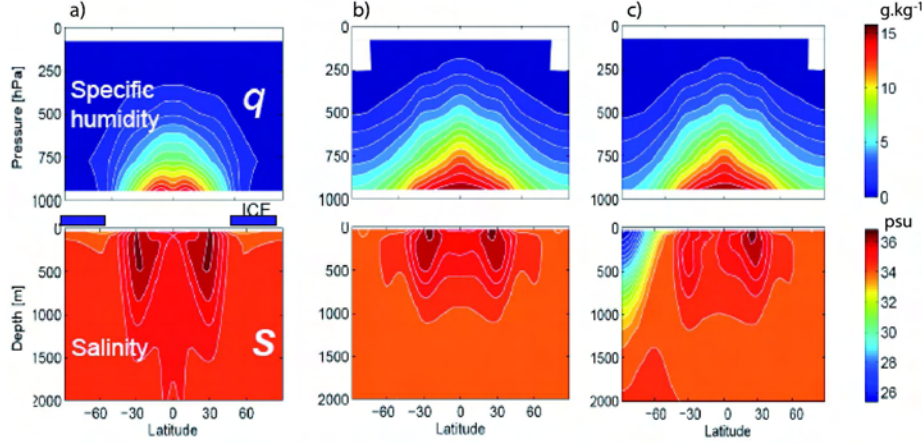


Figure 11: Specific humidity of the atmosphere (top) and salinity of the ocean (bottom) computed in the aquaplanet (a), the ridge (b) and the drake (c) models.

5.2 Heat Transport and its Partition

The ridge enhances the heat transport in ocean towards the poles (figure 12) explaining the absence of ice. In the drake model, there is less heat transport in the southern hemisphere and consequently an ice cap can be sustained over this pole. In the case of a ridge with an equatorial passageway, the heat transport is enhanced leading to warmer poles (figure 13). Indeed, due to less pressure gradients at the equator, winds are balanced by ageostrophic poleward flow.

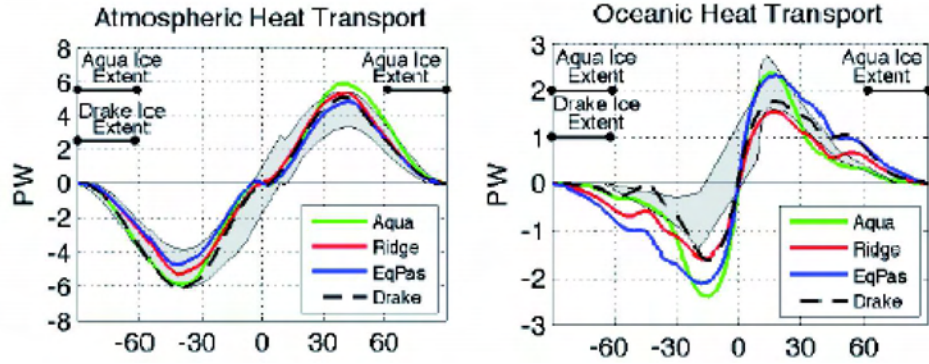


Figure 12: Atmospheric (left) and oceanic (right) heat transport computed with the different models. The results of the ridge model with an equatorial passageway (blue curve) are also shown.

In this section, theoretical considerations of a 2-box model will be presented. A box is shared between two parts: the first part represents the equatorial zone and the second one is the polar area (figure 14). A meridional energy flux f is transferred from the equatorial

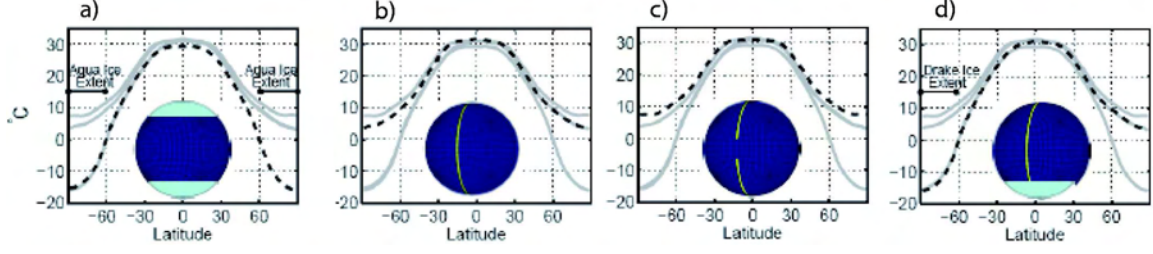


Figure 13: Surface air temperature (dashed line) for the aquaplanet (a), the ridge (b), the ridge with an equatorial passageway (c) and the drake (d) models.

box to the polar box. Each box undergoes an incoming solar radiation s and an outgoing longwave radiation i .

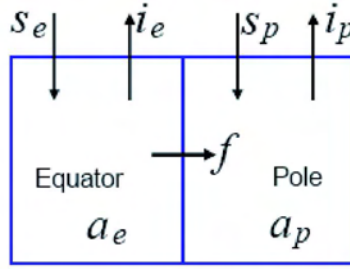


Figure 14: The 2-box model.

The energy balance in the two box is

$$a_e s_e - i_e - f = 0, \quad (4)$$

$$a_p s_p - i_p + f = 0 \quad (5)$$

with a the co-albedo.

Each quantity (s , i or a) can be decomposed as

$$(\cdot)_e = (\bar{\cdot}) + \frac{1}{2}\Delta(\cdot), \quad (6)$$

$$(\cdot)_i = (\bar{\cdot}) - \frac{1}{2}\Delta(\cdot) \quad (7)$$

where $(\bar{\cdot})$ is for the global average of (\cdot) and $\Delta(\cdot) = (\cdot)_e - (\cdot)_i$.

Using equations 4 and 5, f can be expressed as

$$f = \frac{1}{2} \left(\Delta s \bar{a} + \bar{s} \Delta a + \frac{1}{2} \Delta a \Delta s - \Delta i \right). \quad (8)$$

Stone [4] argued that $\bar{s} \Delta a$ is balanced by Δi and that $\frac{1}{2} \Delta a \Delta s$ is negligible (see figure 15). Therefore $\Delta s \bar{a}$ is the predominant term in the meridional energy flux

$$f \simeq \frac{1}{2} \Delta s \bar{a}. \quad (9)$$

It is a reasonable gross estimate but in error by order 20%. This means that the total heat transport only depends on planetary albedo, solar constant and radius of planet. However if ice caps cover the poles, the excess of albedo Δa is important and the $\Delta s \bar{a}$ effects can not be ignored.

	$\Delta s \bar{a}$	$\bar{s} \Delta a$	$\frac{1}{2} \Delta s \Delta a$	$-\Delta i$	f	Implied dimensional heat transport
Aqua	0.240	0.086	-0.016	-0.021	0.289	6.25PW
Ridge	0.249	0.023	-0.004	0.002	0.270	5.88PW
Drake, NH	0.250	0.024	-0.004	0.007	0.277	6.03PW
Drake, SH	0.239	0.087	-0.016	-0.026	0.285	6.21PW

Figure 15: Estimates of the different terms of the calculation of the meridional energy flux f for the different models.

5.3 Double Drake Model

In the double drake model, the ocean is divided into a small and a large basin which can communicate in the southern hemisphere (figure 8d).

The small basin is warmer and more salty than the large one (figure 16). The water is reduced in the small basin by evaporation and precipitates in the large basin. The salt water in the small basin convects and thus a self-sustained overturning circulation is present in the small basin (figure 17). The convection takes place preferentially in the small basin.

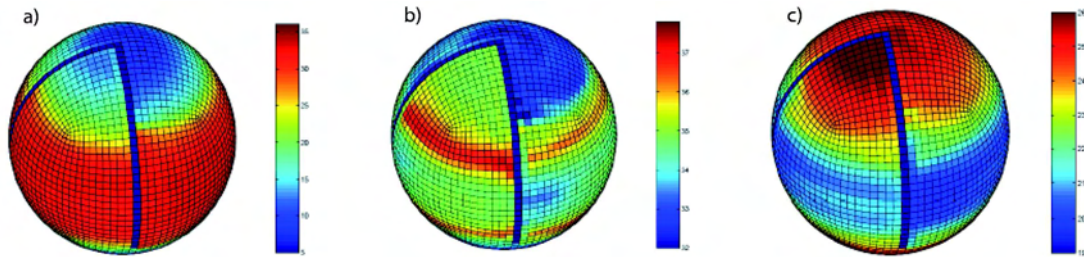


Figure 16: Results of the double drake model: sea surface temperature (a), sea surface salinity (b) and sea surface density (c).

The agreement in heat transport by atmosphere and ocean between today's climate and the double drake model is good despite the straightforward geometrical constraints of the model (figure 18). In particular, the asymmetry in the ocean heat transport between the two hemisphere is well reproduced.

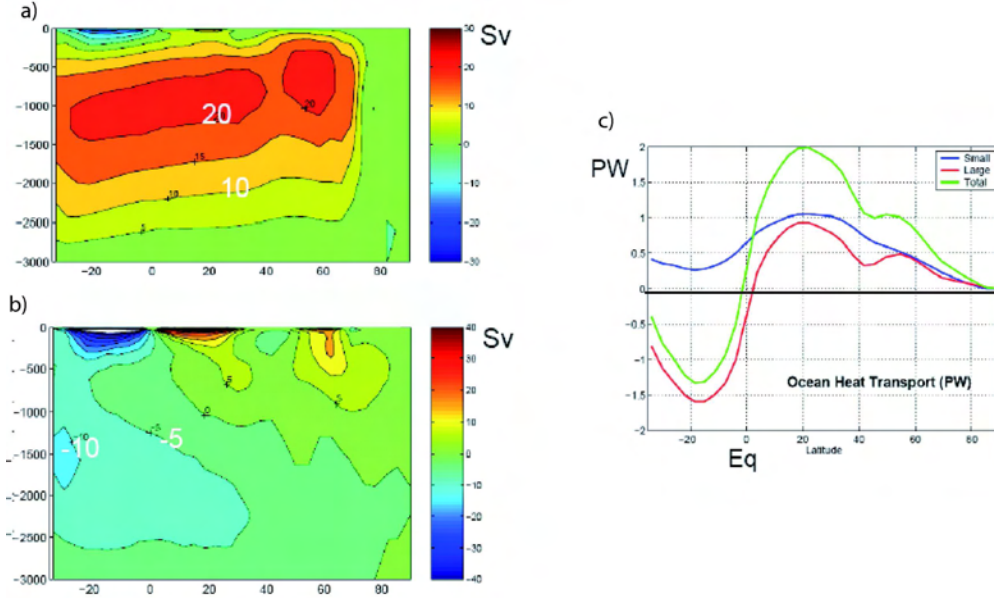


Figure 17: Left: Overturning circulation in the small basin (a) and in the large basin (b) for the double drake model. Right: Ocean heat transport in the two basins (c).

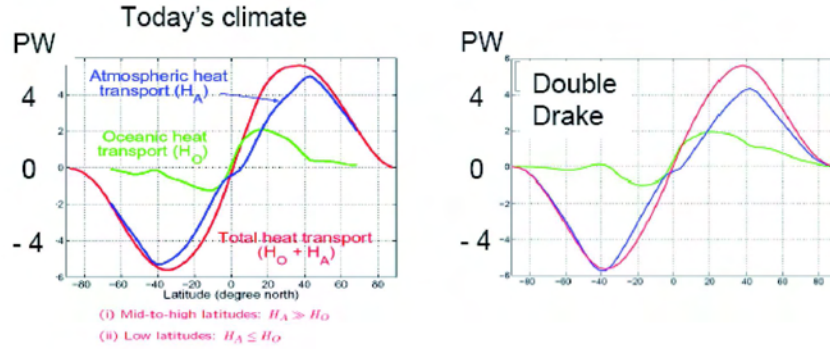


Figure 18: Left: Heat transport in the Earth's current climate. Right: Heat transport in the double drake model.

6 Conclusion

Some points deserve to be emphasized:

- Energy flux partition can be rationalized by

$$\frac{H_A}{H_O} = \frac{\Psi_A}{\Psi_O} \times \frac{\Delta B_A}{\Delta B_O}.$$

The dominance of H_A over H_O is a consequence of $\Psi_A \gg \Psi_O$.

- Although H_O is small, it plays a crucial role *e.g.* in controlling sea ice extent. The climates of the different models introduced above are very different to one-another especially due to the presence/absence of ice at the poles.
- Oceanic energy transport is primarily achieved by wind-driven currents (not buoyancy-driven flow).
- $f \simeq \frac{1}{2}\Delta s\bar{a}$ is a good predictor of total energy transport in warm climate. However $\Delta s\bar{a}$ effects are large in cold climates and can not be ignored.
- Opening of drake passage has a profound effect on global circulation: it is a major source of inter-hemispheric asymmetry.

Some challenges of this work are:

- Multiple equilibria: indeed, stable solutions without ice can be found for the aqua-planet.
- Overlaying biogeochemical cycles in these models.
- Connecting to paleo climate: the different geometrical constraints can be seen as the geological evolution (figure 19).
- Further collaborations : include land.

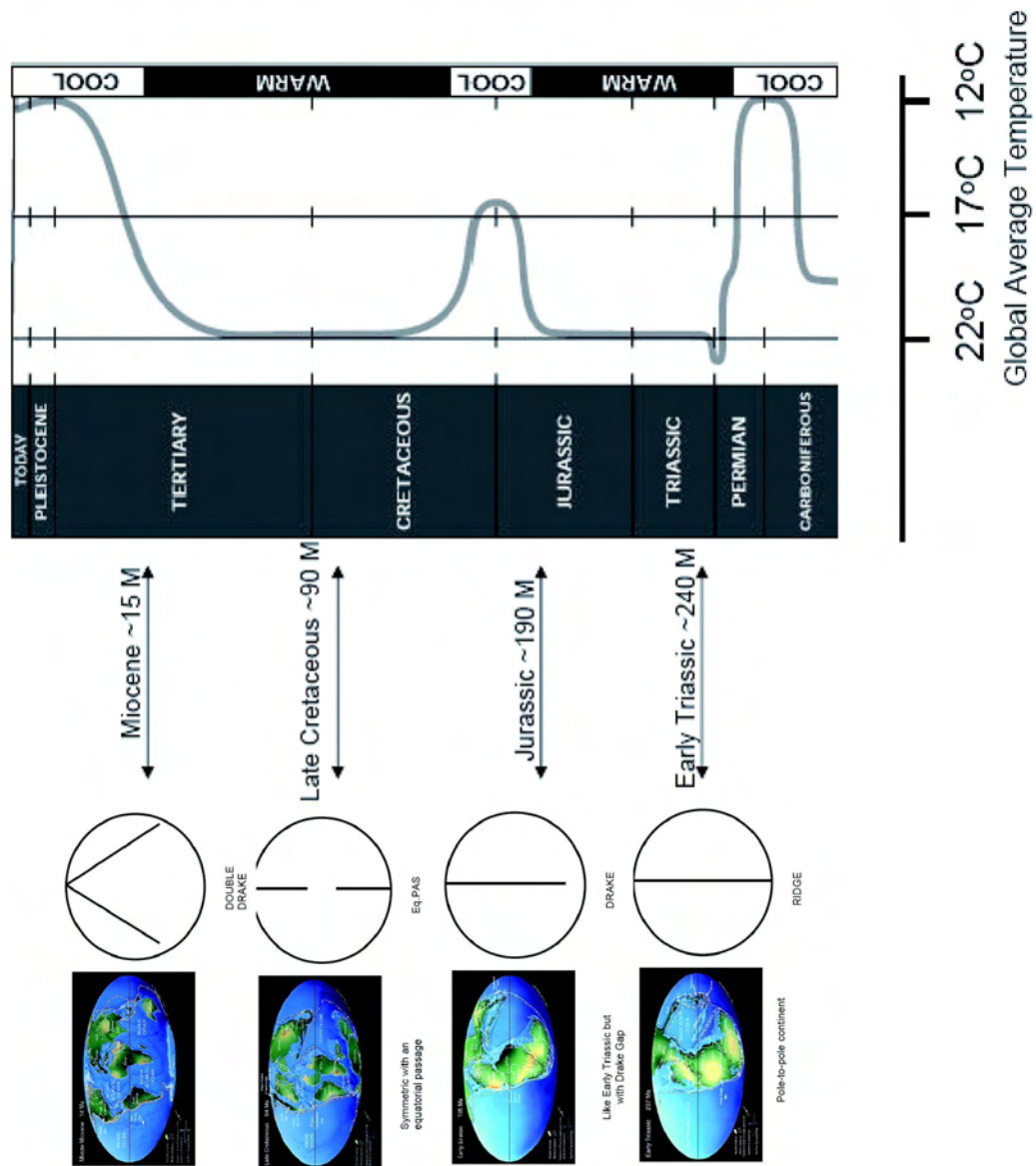


Figure 19: Geological evolution

References

- [1] *MITgcm Online Documentation*.
- [2] *Estimates of meridional atmosphere and ocean heat transports*, Journal of Climate, 14 (2001), pp. 3433–3443.
- [3] R. H. STEWART, *Introduction to physical oceanography*. 2007.
- [4] P. STONE, *Constraints on dynamical transports of energy on a spherical planet*, Dynamics of Atmospheres and Oceans, 2 (1978), pp. 123–139.

Atmospheric Escape

Lecturer: Ray Pierrehumbert

Notes written by: Chris Cawthorn & Yutian Wu

24 June 2008

1 Introduction

The Earth's atmosphere is a comparatively thin layer of a gaseous mixture which is distributed almost uniformly over the surface of the Earth. In the vertical direction, more than 99% of the mass of the atmosphere is found below an altitude of 30km. In contrast, the horizontal scale of the atmosphere is of order 20,000km. The atmosphere is composed of several layers which differ in composition, temperature, stability and energetics. Starting from the surface, the main layers are the troposphere, stratosphere, mesosphere, and thermosphere, separated by conceptual partitions called pauses (e.g., tropopause). The concentrations of nitrogen, oxygen and some inert gases are practically uniform in the atmosphere up to the mesopause. This region constitutes the homosphere. However, above about 100km, the density of gas begins to fall off exponentially with increasing altitude with a rate depending on the molecular mass. Larger mass constituents, such as oxygen and nitrogen, fall off more quickly than lighter ones such as helium and hydrogen. This layer, in which the composition of the atmosphere varies with altitude, is called the heterosphere.

In the following, we introduce atmospheric escape. Due to thermal mechanisms, a lighter molecule is more likely to escape from the atmosphere because of its higher average speed at a given temperature. For example, hydrogen escapes more easily than carbon dioxide. This has numerous applications in astrophysical and planetary science.

1.1 Loss of water from Venus

The atmosphere of Venus contains only 0.1 – 1% H_2O , a fact revealed by the recent Mariner 5 and Venera 4 missions. The total abundance of H_2O is $20 - 200\text{gm cm}^{-2}$ in the Venusian atmosphere compared with $320,000\text{gm cm}^{-2}$ in the Earth's atmosphere. The origin of the present atmosphere of Venus is assumed to be the same as that of the Earth, given the similarity in size and mass of the two planets. However, a large amount of H_2O has been lost during Venus' history. In the atmosphere of Venus, water vapor was able to become a major constituent at a high altitude where the atmospheric cold trap (places where the major constituents of the atmosphere condense) was located, and could be steadily photodissociated. The hydrogen atoms freed as a result of this process then flowed outward

from the planet due in part to hydrodynamic escape. This process could account for the large loss of water from Venus. See [2, 4] for a much more thorough description.

1.2 Hydrogen content of the early Earth atmosphere

Research has shown that hydrogen was one of the major constituents in the ancient atmosphere. In addition, from a biological view, the existence and efficient production of prebiotic organic compounds on early Earth was necessary for the origin of life. H_2 , along with O_2 and CO_2 , can absorb extreme ultraviolet (EUV) radiation, but only H_2 can carry energy back to space by hydrodynamic escape. When hydrogen became the major gas in the heterosphere and the major absorber of EUV, the escape rate of hydrogen could have been controlled by the solar EUV flux available to drive the flow. One could consider that a balance then formed between volcanic hydrogen outgassing and the hydrodynamic escape of hydrogen from the atmosphere, helping to maintain the high hydrogen mixing ratio on early Earth. A discussion is given in [6], some of the details of which are given in §3 of these notes.

1.3 Loss of hydrogen from Titan

Methane gas is abundant in Titan’s atmosphere, but it can be broken apart by ultraviolet light via the following process



subsequently reforming to create ethane (C_2H_6). This reaction is common in the atmospheres of giant planets, where the hydrogen remains in the atmosphere due to the high gravitational energy that must be overcome in order for it to escape. However, on Titan, where the gravitational attraction is much lower, hydrogen can escape, causing the observed carbon to hydrogen ratio, which is higher than for pure methane.

1.4 Stellar wind

The phenomenon of atmospheric escape is not confined merely to planetary masses. Indeed, the stellar wind is an escape process that occurs at the outer limits of the Sun’s atmosphere. Light elements, particularly hydrogen, gain sufficient energy to escape the Sun’s gravity, and are radiated outwards. So great is the magnitude of this release that it can be measured from Earth. The process, while essentially hydrodynamical, is complicated by the influence of the solar magnetic field, and the fact that the high solar atmosphere is a plasma. We shall not discuss the solar wind further in this document, but direct the reader to [5] for a detailed discussion of the subject.

2 Transcritical flows

It is believed that atmospheric flows must be supersonic in order to escape the planet’s gravitational field. This viewpoint is justified, somewhat dubiously, by the assertion that the steady-state outflow of subsonic fluid has an infinite density everywhere, whereas the

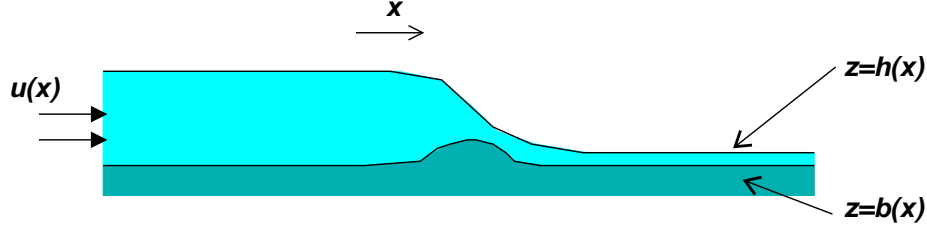


Figure 1: A sketch of shallow-water flow over a gentle bump.

outflow of supersonic fluid can support a more reasonable steady density distribution. Given the unfathomable size of the Universe, however, it seems unlikely that a steady state can be reached, making this argument rather spurious. Much more reasonable is a state in which material gradually radiates away from the planet. However, this transient state cannot be used to draw conclusions about the velocity of escape. Nevertheless, we shall proceed with the belief that flows involved in atmospheric escape must transition between subsonic atmospheric flows and supersonic exospheric flows - that is, they must be *transcritical*. Transcritical flows, characterised by a singularity in the differential equations governing the fluid, are abundant in both modern and classical fluid mechanical literature. In this section, we shall introduce the notion of a transcritical flow via some familiar examples.

2.1 Shallow-water flow over a bump

The flow of a shallow layer of water over a gentle bump is one of the simplest examples of a physical system exhibiting transcritical behaviour, and is often covered in introductory courses in fluid mechanics. The situation is illustrated by Figure 1, with fluid contained between a free surface $z = h(x)$ and a rigid base $z = b(x)$. We denote the fluid velocity (assumed uniform in depth) by $u(x)$. The fluid is incompressible, with constant density ρ .

The equations of mass and momentum conservation applied to the flow, assuming a steady-state solution, are

$$\partial_x ((h - b)u) = \partial_x \Phi = 0 \quad (2)$$

$$u \partial_x u = -g \partial_x h. \quad (3)$$

By differentiating (2) and substituting for $\partial_x h$ in (3), we arrive at the equation

$$u (1 - \text{Fr}^{-2}) \partial_x u = -g \partial_x b, \quad (4)$$

where the Froude number is given by the ratio of the local velocity to the local wave speed, thus

$$\text{Fr} = \frac{u}{\sqrt{g(h - b)}}. \quad (5)$$

If we assume that the basal height $b(x)$ is given, then (4) is an ordinary differential equation for $u(x)$, with forcing $-g \partial_x b$. We note, however, that this equation is singular for $u = 0$ and $\text{Fr} = 1$. The former case is simply the degenerate case of no flow, but the latter is much more important. If at any point the flow conditions are such that $\text{Fr} = 1$, (4) may only have

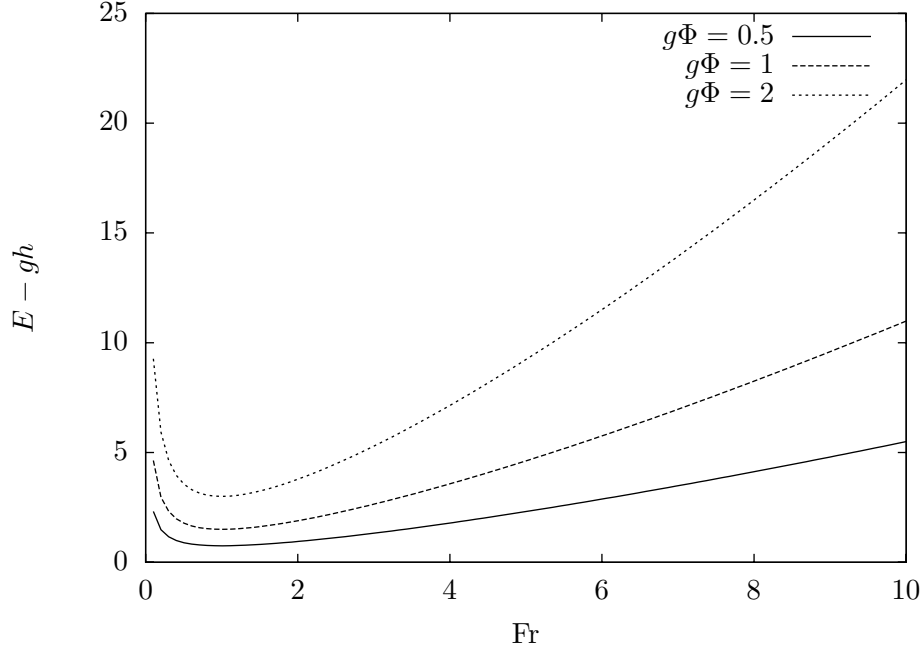


Figure 2: A sketch of the relationship between Froude number and energy, as given by (10)

a smooth solution if the forcing also vanishes at that point - i.e. $\partial_x b = 0$ where $\text{Fr}(x) = 1$. Crucially, this means that the flow may only transition smoothly between a region where $\text{Fr} < 1$ (a *subcritical* region) and a region where $\text{Fr} > 1$ (a *supercritical* region) at a point where $\partial_x b = 0$.

One can think of this transcriticality in the context of energy. If we simply rearrange (3) into the conservative form

$$0 = \frac{\partial}{\partial x} \left(\frac{1}{2} u^2 + gh \right) \quad (6)$$

$$= \frac{\partial}{\partial x} \left(\frac{1}{2} u^2 + g(h - b) + gb \right) \quad (7)$$

$$= \frac{\partial}{\partial x} \left[(h - b) \left(\frac{1}{2} \text{Fr}^2 + 1 + \frac{b}{h - b} \right) \right] \quad (8)$$

$$= \frac{\partial}{\partial x} \left[\left(\frac{g\Phi}{\text{Fr}} \right)^{2/3} \left(\frac{1}{2} \text{Fr} + 1 + \frac{b}{h - b} \right) \right]. \quad (9)$$

If we define the x -independent quantity here to be a specific energy E , then (9) relates E and Fr via

$$E - gh = \left(\frac{g\Phi}{\text{Fr}} \right)^{2/3} \left(\frac{1}{2} \text{Fr}^2 + 1 \right). \quad (10)$$

This relationship is shown in Figure 2. Evaluating E by using the upstream boundary conditions, we find that either zero, one, or two solutions for Fr exist at any location x ,

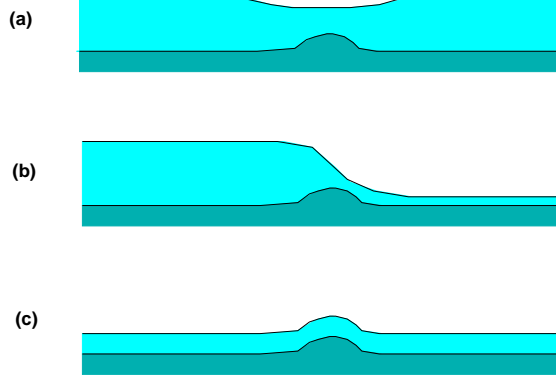


Figure 3: Sketches of the shallow-water flow over a bump for (a) subcritical flow, (b) transcritical flow, and (c) supercritical flow.

given the local bed height $b(x)$. A smooth solution for $\text{Fr}(x)$ must remain on the curve shown in Figure 2, and take the local value of $E - gb$. It should be noted that the minimum energy occurs when $\text{Fr} = 1$, i.e. at the transcritical point. This agrees with our earlier conclusion that, in order to pass smoothly from subcritical to supercritical flow, one must pass through a region where h is extremal. If the geometry and upstream conditions are such that the transcritical regime can be reached, there are three possibilities for the flow, given an initial upstream energy E . These are illustrated in Figure 3.

If the bed elevation $b(x)$ should become sufficiently large so as to forbid any solution of (10), then we can expect no steady solution to exist. In practice, a phenomenon known as ‘choking’ occurs. Fluid builds up behind the bump, creating a disturbance whose upstream extent increases in time until it has sufficiently modified the upstream boundary conditions, allowing a steady flow to form.

2.2 Compressible flow in a duct

Another important example of transcritical behaviour originates as a problem in supersonic propulsion. Consider the duct illustrated in Figure 4. In a simple, one-dimensional model, the compressible fluid contained in the duct has local density $\rho(x)$, and velocity $u(x)$. The cross-sectional area of the duct is denoted by $A(x)$. Once again, we can write down equations describing the conservation of mass and momentum

$$\partial_x (\rho A u) = \partial_x \Phi = 0, \quad (11)$$

$$u \partial_x u = -\frac{1}{\rho} \partial_x p, \quad (12)$$

but in this case we must supplement these equations with an equation of state. Here, it suffices to use the general form for an adiabatic gas

$$p = p(\rho), \quad \text{with sound speed} \quad c_s^2 = \frac{dp}{d\rho}. \quad (13)$$

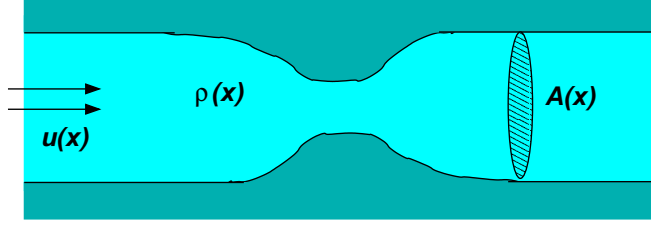


Figure 4: Diagram and notation for compressible duct flow

We can now use (13) to substitute ρ for p in (12), before using conservation of mass to eliminate $\partial_x \rho$, as we did in §2.1. This quickly leads to the ordinary differential equation

$$\frac{1}{u} (\mathcal{M}^2 - 1) \partial_x u = \frac{\partial_x A}{A}, \quad (14)$$

where $\mathcal{M} = u/c_s$ is the *Mach number*. Once again, for a given geometry defined by $A(x)$, the steady-state velocity obeys a singular ordinary differential equation. Importantly, the singular point occurs when $\mathcal{M} = 1$, separating subsonic (subcritical) flow from supersonic (supercritical) flow. In order to have a transsonic flow, passing smoothly through $\mathcal{M} = 1$ (known as the *Mach point*) we require that the forcing in (14) must vanish. That is, the area $A(x)$ must be extremal at the Mach point. This is known in aerodynamics as the *minimum area rule for transsonic flow*.

Framing this problem in terms of energy requires use of the thermodynamical relation can proceed by finding a conservative form in much the same way as in §2.1

$$0 = \frac{\partial}{\partial x} \left[\frac{1}{2} u^2 + \int^\rho \frac{c_s^2(\rho')}{\rho'} d\rho' \right] \quad (15)$$

$$= \frac{\partial}{\partial x} \left[c_s^2 \left(\frac{1}{2} \mathcal{M}^2 + \frac{1}{c_s^2} \int^\rho \frac{c_s^2(\rho')}{\rho'} d\rho' \right) \right]. \quad (16)$$

One would now like to define a specific energy $E(\mathcal{M})$ to be equal to this conserved quantity, but the dependence of c_s on ρ means that this is as far as a general adiabatic theory can take us. If we consider specifically an ideal gas, where $p = p_0(\rho/\rho_0)^{1/\beta}$, then we may write (16) in the form

$$0 = \frac{\partial}{\partial x} \left[\left(\frac{p_0 \rho^{-1+1/\beta}}{\beta \rho_0^{1/\beta}} \right) \left(\frac{1}{2} \mathcal{M}^2 + \frac{\beta}{1-\beta} \right) \right]. \quad (17)$$

We can then use conservation of mass (11) to substitute for ρ and arrive at the equation

$$0 = \frac{\partial E}{\partial x} = \frac{\partial}{\partial x} \left[c_{s0}^2 \left(\frac{\mathcal{M}_0 A_0}{\mathcal{M} A} \right)^{2 \frac{1-\beta}{1+\beta}} \left(\frac{1}{2} \mathcal{M}^2 + \frac{\beta}{1-\beta} \right) \right], \quad (18)$$

where the reference values $(c_{s0}, \mathcal{M}_0, A_0)$ are evaluated using the upstream boundary conditions. Note now that the specific energy E is written only as a function of the local Mach number \mathcal{M} , given the variation of cross-sectional area A with x . This relationship, for a representative value of β , is illustrated in Figure 5

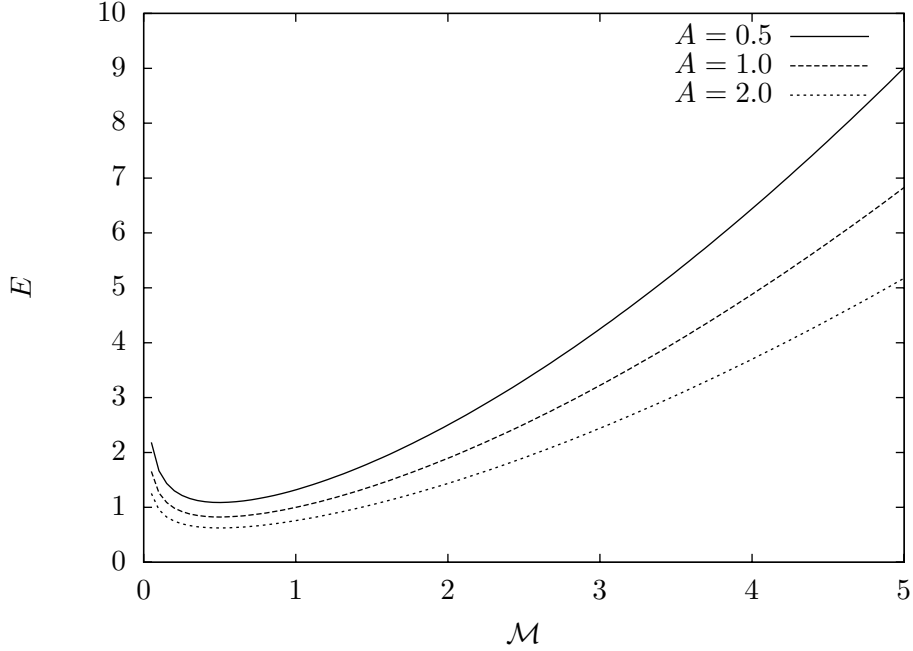


Figure 5: A plot of the relationship between energy and Mach number, as given by (18), with $\beta = 2/3$.

Many similarities exist between this and the flow over a bump discussed in §2.1. The principal difference is that the varying geometry, specified by $A(x)$, appears as a multiplicative modification to the energy E , rather than as an additive modification in the shallow bump problem. This changes the behaviour surprisingly little. At a given point, there are still zero, one, or two solutions, with one sub- and one super-critical in the latter case. If the cross-sectional area decreases sufficiently, it can be that no steady solution exists, in which case the ‘choking’ effect causes an upstream disturbance that modifies the upstream boundary conditions.

2.3 Other examples

There are many other examples of transcritical behaviour occurring in many areas of fluid mechanics and physics in general. For example, calculations regarding the stellar wind result in a singular differential equation, forcing flow to pass through a critical point - the *Alfvén point* of magnetohydrodynamics.

An important geophysical example is the flow of a stratified fluid over an mountain. If the atmosphere could be assumed to have a rigid lid, then a weakly nonlinear analysis would proceed analogously to the shallow water flow discussed in §2.1. The same is not true in an unbounded region, but the phenomenology is similar, as shown by Pierrehumbert and Wyman [3]. Figure 6 illustrates their computational solutions, which show that the flow is subcritical upwind of a mountain, but supercritical on the lee-side in a typical geophysically-

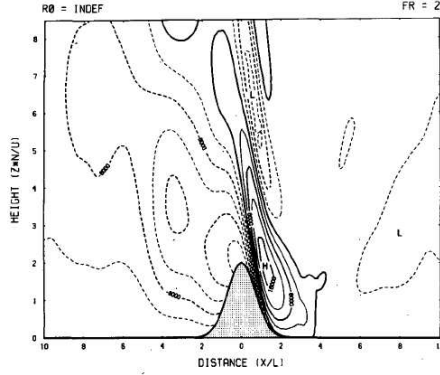


Figure 6: Streamlines of perturbation to uniform left-to-right flow past an idealized mountain, from [3].

reasonable region of parameter space. In their work, Pierrehumbert and Wyman [3] found that the dimensionless Froude number controlling the flow is given by

$$\text{Fr} = \frac{U}{NH_m}, \quad (19)$$

where U is the local velocity scale, N the buoyancy frequency, and H_m a vertical length-scale comparable to the height of the mountain. For small values of this parameter, a ‘blocking’ phenomenon occurs. Similar to the ‘choking’ behaviour described in the earlier examples, blocking generates an upstream-propagating columnar mode, which has the effect of modifying the upstream boundary conditions by creating a layer of stagnant fluid near the ground, effectively reducing the height of the mountain and allowing a steady flow to be reached.

3 Atmospheric escape

We have previously alluded to the fact that atmospheric escape flows are complicated by a number of factors. Solar radiation, gravitational potential and non-adiabatic effects all come into play, among many other effects. A brief discussion of the control of hydrogen in the early Earth atmosphere (see [6]) makes mention of several complications. In this section, we shall consider a simple, spherically-symmetric model for atmospheric escape, before mentioning some possible refinements and future directions for interesting research.

3.1 A spherically-symmetric model

In order to formulate a spherically-symmetric model for atmospheric escape, we follow a very similar approach to the compressible flow problem described in §2.2. In this case, conservation of mass may be expressed in the form

$$\Phi = 4\pi r^2 \rho(r) u(r) = \text{const.} \quad (20)$$

Note that, in this case, there is no extremal ‘cross-sectional area’, equivalent to $A(x)$ in the case of the duct. We must therefore rely on a different mechanism to lead to transcritical behaviour. When considering the conservation of momentum, we must therefore include the gravitational potential $\chi = -GM/r$, where M is the mass of the planet, and G is the gravitational constant. Assuming an adiabatic gas, as in §2.2, we can arrive at the following form of the momentum equation

$$\frac{1}{u} (\mathcal{M}^2 - 1) \frac{\partial u}{\partial r} = \frac{1}{(c_s r)^2} (2rc_s^2 - GM). \quad (21)$$

For this formulation, it is clear that a singular point exists when $\mathcal{M} = 1$. At this point, one may only transition smoothly from a subsonic flow to a supersonic flow provided that

$$c_s(R_s) = \sqrt{\frac{GM}{2R_s}}, \quad (22)$$

where R_s is the *sonic radius*, also referred to as the *sonic point*. In this case, it is the gravitational potential term that plays the role of the bump in §2.1 or the constriction of the duct in §2.2.

Arranging the momentum equation into conservative form yields the equation

$$0 = \frac{\partial E}{\partial r} = \frac{\partial}{\partial r} \left[\frac{1}{2} u^2 + \int^\rho \frac{c_s^2(\rho')}{\rho'} d\rho' - \frac{GM}{r} \right], \quad (23)$$

which allows us to define a specific energy E . For an ideal gas, as described in §2.2, it is possible to write

$$E = c_{s0}^2 \left(\frac{\mathcal{M}_0 A_0}{\mathcal{M} A} \right)^{2\frac{1-\beta}{1+\beta}} \left(\frac{1}{2} \mathcal{M}^2 + \frac{\beta}{1-\beta} \right) - g_0 R_0 \left(\frac{R_0}{r} \right), \quad (24)$$

where the terms represent specific kinetic, internal and potential energy, respectively. In the case of atmospheric escape, the temperatures are not usually large enough for internal heat to be a significant component of energy, so the principal balance is between kinetic and potential energies. This leads, as every good schoolchild knows, to the notion of an *escape velocity* $u_e = 2g_0 R_0$, at which a particle has enough kinetic energy to escape the potential well of the planet. This balance is sufficient, in order-of-magnitude, to drive an outflow like the hydrogen-rich solar wind. However, for typical conditions in the oxygen-rich high terrestrial atmosphere, escape requires an additional input of energy.

3.2 Extensions and generalisations

In real planetary situations, such as that of the Earth, there is little reason to believe that any escape flow will be spherically-symmetric. Indeed, given the difference in solar radiation absorbed on the day- and night-sides of a planet, an entirely symmetric flow would be very surprising. It is possible that the heating the sunward side of the atmosphere by EUV radiation drives a flow to the dark side of the planet, where it is then able to escape. The idea is that the radiatively-driven circulation forces the exospheric material into a jet directed away from the Sun. Calculations based on this notion have yet to be carried out.

Furthermore, our assumption that the gas is adiabatic is not valid for the case of atmospheric escape from Earth’s atmosphere. This is clear from the calculations, such as that of §3.1, which show that there is insufficient energy for atmospheric escape of oxygen. It is therefore necessary to include some non-adiabatic effects, such as local heating by EUV radiation, in order to arrive at correct scalings for terrestrial atmospheric escape. It should be noted, however, that simple theories of adiabatic hydrodynamic escape perform admirably when applied to the solar wind.

An interesting model through which to consider non-adiabatic effects in hydrodynamical escape might be to return to transsonic flow in a duct, but to supply heat, uniformly or otherwise, to the gas in the duct. This heating should weaken the conditions leading to the minimum area rule, perhaps displacing the sonic point, or allowing for multiple sonic points. When generalised to atmospheric collapse, this could result in enough energy being supplied to allow material to escape.

It should be noted that hydrodynamical escape is not the only mechanism allowing material to leave Earth’s atmosphere. The article by Catling [1] argues that hydrodynamic escape may not even be the dominant mechanism for hydrogen removal, and that so-called ‘non-thermal’ escape mechanisms are more significant. One hopes that a better understanding of the details of the various mechanisms may lead to a resolution of the argument, and a clearer picture of how planetary atmospheres can evolve.

References

- [1] D. C. CATLING, *Comment on “a hydrogen-rich early earth atmosphere”*, Science, 311 (2005), p. 38a.
- [2] J. F. KASTING AND J. B. POLLACK, *Loss of water from venus. i. hydrodynamic escape of hydrogen*, Icarus, 53 (1983), pp. 479–508.
- [3] R. T. PIERREHUMBERT AND B. WYMAN, *Upstream effects of mesoscale mountains*, J. Atmos. Sci., 42 (1985), pp. 977–1003.
- [4] S. I. RASOOL, *Loss of water from venus*, J. Atmos. Sci., 25 (1968), pp. 663–664.
- [5] M. J. THOMPSON, *An introduction to astrophysical fluid dynamics*, Imperial College Press, 2006.
- [6] F. TIAN, O. TOON, A. A. PAVLOV, AND H. DE STERCK, *A hydrogen-rich early earth atmosphere*, Science, 308 (2005), p. 1014.

Double-Diffusive Convection

Lecturer: Timour Radko

Notes written by: Christophe Gissinger & Toby Wood
Woods Hole Geophysical Fluid Dynamics Program 2008

27 June 2008

1 Introduction

Even in systems with a negative vertical density gradient, instability is possible if the density is controlled by two components (e.g. temperature T and salinity S) that diffuse at different rates [24]. In the oceans, temperature diffuses approximately 100 times faster than salt, and so many regions of the ocean are potential candidates for so-called “double-diffusive instability”.

Doubly-diffusive effects were observed by several authors before Stern [24] explained the physical mechanism responsible. Jevons [4] and Ekman [1] had previously observed instability at the interface between temperature stratified water and an overlying layer of denser fluid, but neither recognised the significance of double-diffusion. Later, Stommel *et al.* [28] showed that a “perpetual salt fountain” can arise when a tube is inserted vertically through the interface between a layer of warm, salty water overlying cold, fresh water. The fountain persists until the system becomes well mixed. This experiment has recently been realised on an industrial scale [29].

There are two forms of double-diffusive instability, referred to as “salt fingers” and “diffusive convection”.

1.1 Salt fingers

Suppose a fluid system contains positive vertical gradients of both temperature T and salinity S , such that the total density gradient is negative (i.e. stably stratified). We describe this configuration as “warm salty over cold fresh” (see Figure 1). In the absence of diffusive effects, a parcel of fluid displaced vertically downwards would find itself more buoyant than its surroundings, and therefore rise. However, if temperature diffusion is sufficiently strong, and salt diffusion sufficiently weak, then the parcel can come into thermal equilibrium with its surroundings before it rises, whilst still remaining salty. It therefore becomes less buoyant than its surroundings, and continues to fall under gravity. This process leads to “salt fingers” propagating down through the medium and, conversely, fresh fingers propagating upward (see Figure 2).

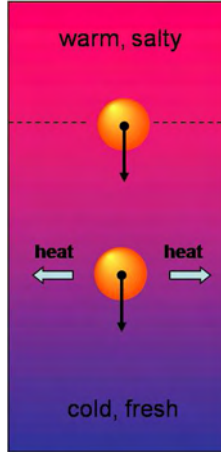


Figure 1: Schematic of the salt finger mechanism.

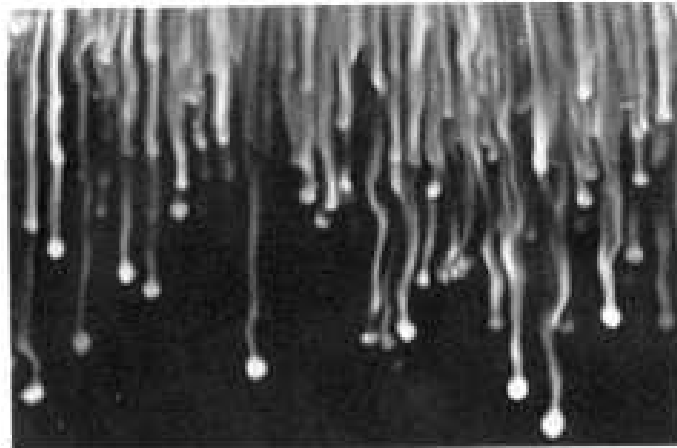


Figure 2: Salt fingers in a laboratory experiment [3].

1.2 Diffusive convection

Suppose we have negative vertical gradients of both T and S such that the total density gradient remains stably stratified, i.e. “cold fresh over warm salty” (see Figure 3). A parcel of fluid displaced vertically downwards now quickly absorbs heat by diffusion from its surroundings, and so “overshoots” when rising back through the medium. This leads to an oscillation of growing amplitude, i.e. an oscillatory instability.

Both kinds of double-diffusive instability are thought to play a role in small-scale mixing of the world’s oceans. However, since the conditions necessary for salt fingers are more common, we focus here primarily on this instability.

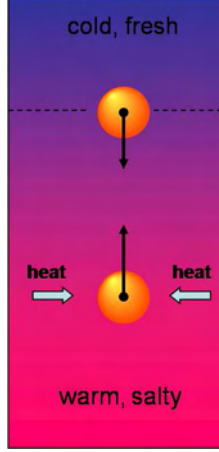


Figure 3: Schematic of the diffusive convection mechanism.

2 Mathematical formulation

2.1 The linear instability

We begin with the Boussinesq equations, which are valid over height scales less than the pressure and density scale heights, provided the flow speed remains much less than the sound speed.

$$0 = \nabla \cdot \mathbf{u}, \quad (1)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{1}{\rho_0} \nabla p - g \frac{\rho}{\rho_0} \hat{\mathbf{k}} + \nu \nabla^2 \mathbf{u}, \quad (2)$$

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = k_T \nabla^2 T, \quad (3)$$

$$\frac{\partial S}{\partial t} + \mathbf{u} \cdot \nabla S = k_S \nabla^2 S, \quad (4)$$

$$\frac{\rho - \rho_0}{\rho_0} = \beta(S - S_0) - \alpha(T - T_0). \quad (5)$$

Here, k_T and k_S represent diffusivity of heat and salt, ν is the kinematic viscosity and (ρ_0, p_0, T_0, S_0) denote the (constant) reference values of density, pressure, temperature and salinity.

We consider linear perturbations to a background state with $\bar{\mathbf{u}} = 0$, $\bar{T}_z \equiv \frac{d\bar{T}(z)}{dz} = \text{const} > 0$, $\bar{S}_z \equiv \frac{d\bar{S}(z)}{dz} = \text{const} > 0$ and $\frac{\partial \bar{p}}{\partial z} = -g\bar{\rho}$. (Overbars are used to denote unperturbed quantities, and primes will be used for their linear perturbations.) The background is stably stratified provided that density is decreasing with height, i.e.

$$\frac{1}{\rho_0} \bar{\rho}_z = \beta \bar{S}_z - \alpha \bar{T}_z < 0. \quad (6)$$

The perturbations then obey the equations

$$0 = \nabla \cdot \mathbf{u}', \quad (7)$$

$$\frac{\partial \mathbf{u}}{\partial t} = -\frac{1}{\rho_0} \nabla p' - g \frac{\rho'}{\rho_0} \hat{\mathbf{k}} + \nu \nabla^2 \mathbf{u}', \quad (8)$$

$$\frac{\partial T'}{\partial t} + w' \bar{T}_z = k_T \nabla^2 T', \quad (9)$$

$$\frac{\partial S'}{\partial t} + w' \bar{S}_z = k_S \nabla^2 S', \quad (10)$$

$$\frac{\rho'}{\rho_0} = \beta S' - \alpha T'. \quad (11)$$

Salt fingers typically exhibit much longer vertical scales than horizontal scales, so we seek solutions that are proportional to $\exp(\lambda t) \sin(k_1 x) \sin(k_2 y)$. We then obtain three algebraic equations relating λ to $\kappa^2 \equiv k_1^2 + k_2^2$:

$$\lambda w' = -g(\beta S' - \alpha T') - \nu \kappa^2 w', \quad (12)$$

$$\lambda T' + w' \bar{T}_z = -k_T \kappa^2 T', \quad (13)$$

$$\lambda S' + w' \bar{S}_z = -k_S \kappa^2 S'. \quad (14)$$

$$(15)$$

For given κ , there are three solutions for λ , one real and two complex conjugate. The complex solutions represent slowly decaying internal waves. The real solution represents double-diffusive instability if $\lambda > 0$. To determine the instability criterion, we analyse the marginal stability condition, $\lambda = 0$; we then find

$$g \left(\frac{\beta \bar{S}_z}{k_S} - \frac{\alpha \bar{T}_z}{k_T} \right) = \nu \kappa^4. \quad (16)$$

This equation has solutions with $\kappa \in \mathbb{R}$ provided that the LHS is positive. The necessary condition for instability is therefore

$$R_\rho \equiv \frac{\alpha \bar{T}_z}{\beta \bar{S}_z} < \frac{k_T}{k_S}. \quad (17)$$

The quantity R_ρ is called the *density ratio*. In the ocean, the diffusivities of temperature and salt are, respectively, $1.4 \times 10^{-7} \text{m}^2 \text{s}^{-1}$ and $1.1 \times 10^{-9} \text{m}^2 \text{s}^{-1}$, so instability requires $R_\rho \lesssim 100$. We also require $R_\rho > 1$ so that the background state is stably stratified (see equation (6)). The growth rate of the salt finger instability is greatest when $R_\rho \approx 1$.

2.2 Instability scales

We assume R_ρ is of order unity, in which case equation (16) implies

$$g \frac{\beta \bar{S}_z}{k_S} \sim g \frac{\alpha \bar{T}_z}{k_T} \sim \nu \kappa^4. \quad (18)$$

The typical horizontal scale of the salt fingers is then d , where

$$d \sim \frac{2\pi}{\kappa} \sim 2\pi \left(\frac{k_T \nu}{g \alpha \overline{T}_z} \right)^{1/4} \quad (19)$$

For typical parameter values ($\alpha \sim 10^{-4} \text{K}^{-1}$, $g \sim 10 \text{m s}^{-2}$, $\overline{T}_z \sim 0.01 \text{K m}^{-1}$, $\nu \sim 10^{-6} \text{m}^2 \text{s}^{-1}$) we find $d \sim 0.06 \text{m}$. The characteristic timescale and velocity for temperature diffusion are then $t \sim d^2/k_T \sim 10^4 \text{s}$ and $v \sim k_T/d \sim 10^{-5} \text{m s}^{-1}$.

3 The role of salt fingers in the oceans

The necessary conditions for double-diffusive instability are common within the world's oceans (see Figure 4). In particular, 90% of the main Atlantic thermocline has density

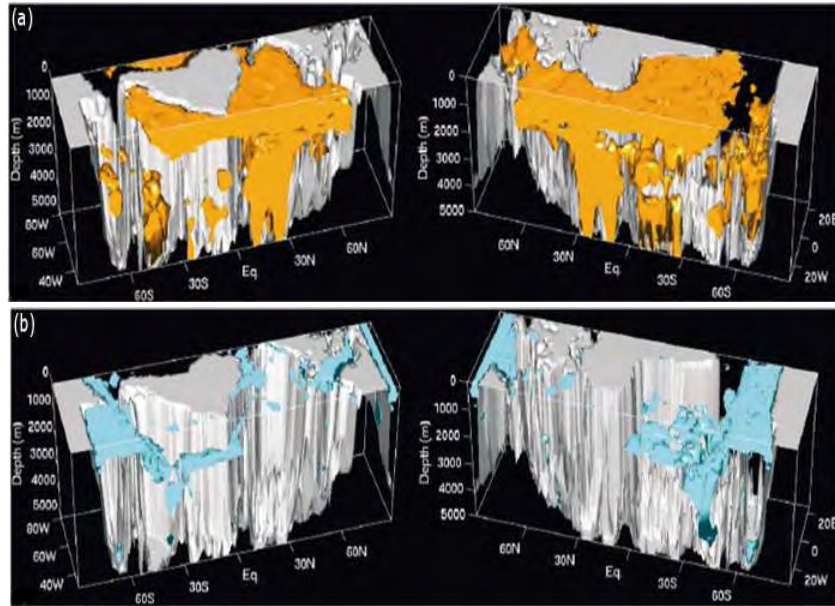


Figure 4: Regions of the ocean potentially susceptible to (a) salt fingers; (b) diffusive convection [35].

ratio $R_\rho < 2.3$, and is therefore strongly unstable to salt fingering. Shadowgraph imaging of the North Atlantic reveals finger-like structures with the centimetre horizontal scale characteristic of salt finger instability (see Figure 5).

Both salt fingers and diffusive convection are readily observed in laboratory experiments and numerical simulations. Importantly, salt fingers are found to be robust to interactions with internal gravity waves at Richardson numbers > 0.5 . At smaller Richardson numbers, horizontal shear interrupts the vertical flux of heat and salt [27].

The main interest in salt fingers within the oceanographic community arises from their implications for vertical mixing [9, 22]. However, they may also explain the temperature–salinity patterns within the oceans (see Figure 6), and in particular the observed “thermohaline staircases” [26, 21]. Finally, salt fingers have been used to explain the lateral

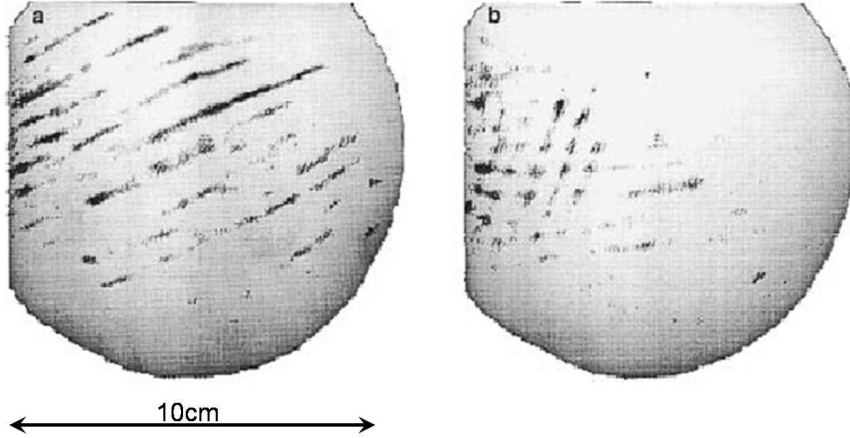


Figure 5: Shadowgraph images of the North Atlantic at 300m depth. The observed filaments are believed to be salt fingers, tilted by shear [9].

intrusions that are regularly observed in the ocean’s temperature and salinity profiles [25].

4 The layered and smooth-gradient regimes of double-diffusive instability

Depending on the density ratio R_ρ , there are two different regimes of salt fingering. For larger values of R_ρ , both T and S have approximately linear vertical gradients, with small perturbations on the scale of the salt fingers. For smaller values of R_ρ , thermohaline staircases develop, with horizontal layers of approximately uniform temperature and salinity separated by thin interfaces (see Figure 7). The vertical scale of these layers is typically much larger than the scale of the fingers themselves, e.g. tens of metres in the ocean, and the transition between these two regimes occurs at around $R_\rho = 1.9$ [21].

Layers also form in numerical and laboratory experiments of oscillatory diffusive convection (see Figure 8). The layers merge until, perhaps, reaching a maximum size. Regarding the interfaces between layers as approximately laminar and steady, the (upward) fluxes of temperature and salt across an interface of thickness h are

$$F_T \sim k_T \frac{\Delta T}{h} \quad \text{and} \quad F_S \sim k_S \frac{\Delta S}{h}, \quad (20)$$

where ΔT and ΔS represent the total variations of temperature and salinity across the interface. The interface is stable provided that the total (upward) density flux, F_ρ , is negative, where $F_\rho = \rho_0(\beta F_S - \alpha F_T)$. We must therefore have

$$\frac{\beta \Delta S}{\alpha \Delta T} < \frac{k_T}{k_S}. \quad (21)$$

We would like to know what determines the vertical flux of T and S through the system and, in particular, how this depends on the density ratio R_ρ . We analyse the staircase and smooth-gradient regimes separately, seeking flux laws in either case.

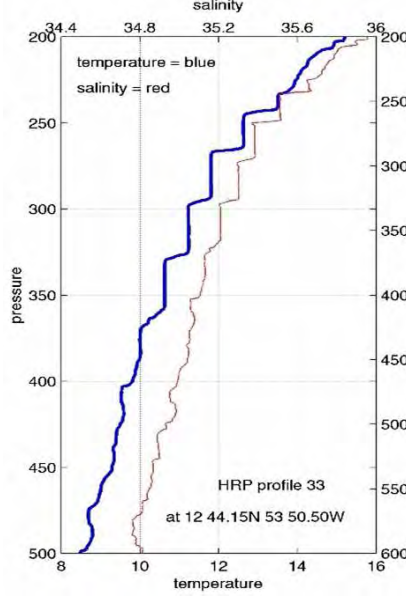


Figure 6: Temperature and salinity profiles in the West Atlantic near Venezuela.

4.1 The interfacial flux laws

We suppose that the fluxes of temperature and salt are determined by conditions within the interfaces, and independent of the height of the layers. By dimensional analysis, we conclude that the dimensionless temperature flux $\frac{F_T/(\Delta T)^{4/3}}{(g\alpha k_T)^{1/3}}$ can depend only on the dimensionless quantities $\frac{\alpha\Delta T}{\beta\Delta S}$, $\frac{\nu}{k_T}$ and $\frac{k_T}{k_S}$. This leads to Turner’s 4/3 flux law [30]:

$$F_T, F_S \propto (\alpha\Delta T)^{4/3}. \quad (22)$$

Laboratory experiments typically exhibit an exponent slightly less than 4/3, though it is unclear to what extent these experiments can be extrapolated to describe the ocean [20, 10, 6, 7]. The hypothesis that temperature and salinity fluxes in the ocean are independent of layer thickness has been tested by Wilson [34]; if the ocean’s thermohaline staircases are in a steady state, then the fluxes, and hence the temperature variations, should be the same across each interface. Figure 9 shows the variation of $\alpha\Delta T$ and $\beta\Delta S$ with layer thickness for a typical Ice-Tethered Profiler (ITP) observation in the Arctic. The observation supports Turner’s key assumption that temperature and salt fluxes are not correlated with the thickness of the layers.

4.2 The gradient flux laws

In the smooth-gradient regime, we are concerned with perturbations to a background state that has linear profiles of T and S . After non-dimensionalising with respect to the scales

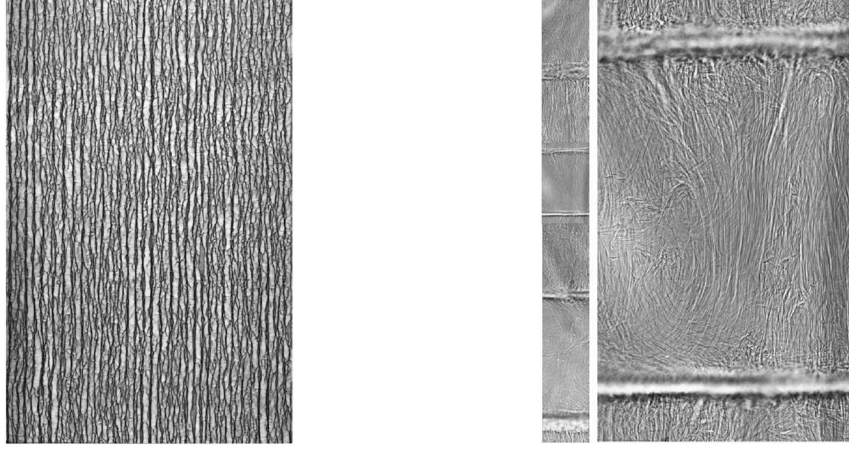


Figure 7: The gradient and staircase regimes of salt fingers [7].

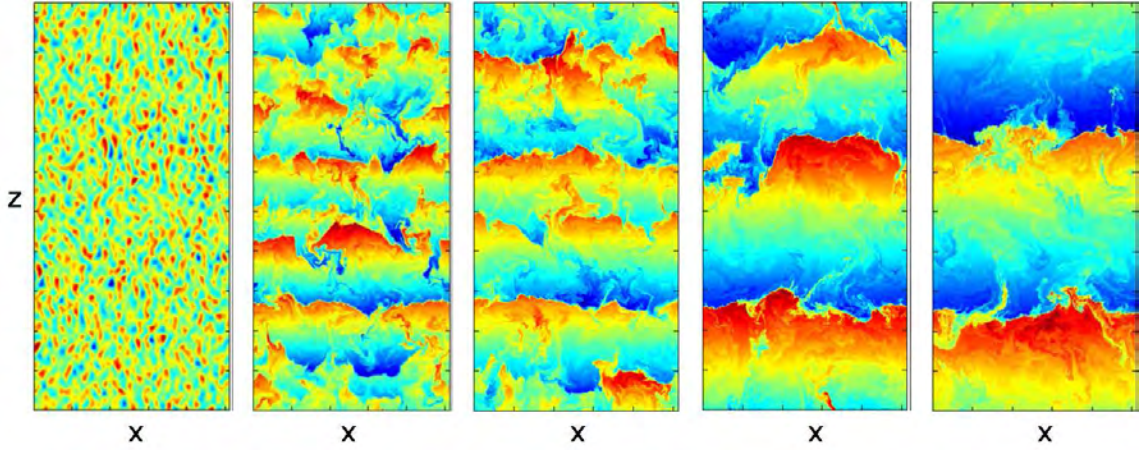


Figure 8: The formation of layers in diffusive convection [12]. The colour shows the departure of temperature from a uniform gradient.

obtained for the linear perturbations, we are left with the equations

$$0 = \nabla \cdot \mathbf{u}, \quad (23)$$

$$\frac{1}{\text{Pr}} \left[\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right] = -\nabla p + (T - S)\hat{\mathbf{k}} + \nabla^2 \mathbf{u}, \quad (24)$$

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T + w = \nabla^2 T, \quad (25)$$

$$\frac{\partial S}{\partial t} + \mathbf{u} \cdot \nabla S + \frac{w}{R_\rho} = \tau \nabla^2 S, \quad (26)$$

where $\tau = k_S/k_T$ is the Lewis number and $\text{Pr} = \nu/k_T$ is the Prandtl number. From these equations, we conclude that the dimensionless temperature and salt fluxes can depend only

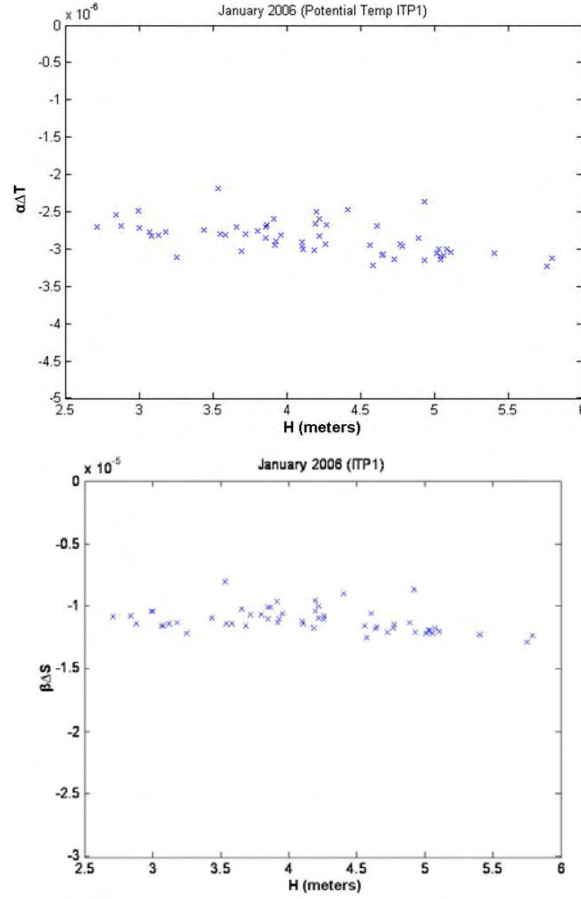


Figure 9: Temperature and salinity changes across interfaces between layers of thickness H .

on the dimensionless numbers R_ρ , τ and Pr . In particular, we can write

$$\frac{-F_T}{k_T \bar{T}_z} = \text{Nu}(R_\rho, \tau, \text{Pr}), \quad (27)$$

where Nu is the Nusselt number. Regarding the diffusivities as fixed, we would like to determine the variation of the Nusselt number with the density ratio R_ρ . Stern [26] noted that salt fingers can excite gravity waves and conjectured that this would lead to a collective instability of the salt finger regime when the Stern number $A = \frac{\beta F_S - \alpha F_T}{\nu(\alpha \bar{T}_z - \beta \bar{S}_z)}$ becomes order unity, in which case the nonlinear regime would be characterised by $A \approx 1$. However, the results of laboratory experiments suggest that the Stern number varies with Prandtl number. Observations of heat-salt fingers typically yield $A = O(1)$, $\text{Pr} \approx 7$ and $\tau \approx 0.01$, whereas sugar-salt fingers have $A \sim 0.001 - 0.1$, $\text{Pr} \approx 1000$ and $\tau \approx 1/3$ [23, 8, 2, 7].

5 Asymptotic analysis of the fingering regime

5.1 Sugar–Salt fingers

Within the high-Prandtl number regime relevant to sugar–salt fingers we can justify neglecting the inertial terms in equation (24). The problem is also simplified by the narrow range of unstable density ratios ($1 < R_\rho < \tau^{-1} \approx 3$). Several authors [31, 32, 13, 5, 14, 18] have undertaken a weakly nonlinear analysis of the marginally unstable modes. The following model is based on an expansion in which $\epsilon = \left(\frac{1}{\tau R_\rho} - 1\right)$ is small [19]. The space and time scales arise from the linear theory, and so we define $\mathbf{x} = \epsilon^{-1/4} \mathbf{x}_0$ and $t = \epsilon^{-3/2} t_0$. The remaining quantities scale as

$$\begin{aligned} \mathbf{u} &\sim \mathbf{x}/t \sim \epsilon^{5/4} \\ \nabla^2 T &\sim \nabla^2 S \sim w \sim \epsilon^{5/4} \\ \Rightarrow (T, S) &\sim \epsilon^{3/4} \\ \nabla p &\sim \nabla^2 \mathbf{u} \\ \Rightarrow p &\sim \epsilon^{3/2} \end{aligned}$$

After expanding the equations in powers of ϵ , the leading order balances are

$$T_0 = S_0 \quad (28)$$

$$w_0 = \nabla_0^2 T_0 = \nabla_0^2 S_0 \quad (29)$$

and the second order balances are

$$\nabla_0 p_0 = (T_1 - S_1) \hat{\mathbf{k}} + \nabla_0^2 \mathbf{u}_0 \quad (30)$$

$$\frac{\partial T_0}{\partial t_0} + \nabla_0 \cdot (\mathbf{u}_0 T_0) + w_1 = \cancel{2\nabla_0 \cdot \nabla_1 T_0} + \nabla_0^2 T_1 \quad (31)$$

$$\frac{1}{\tau} \left[\frac{\partial S_0}{\partial t_0} + \nabla_0 \cdot (\mathbf{u}_0 S_0) \right] + w_0 + w_1 = \cancel{2\nabla_0 \cdot \nabla_1 S_0} + \nabla_0^2 S_1 \quad (32)$$

$$(33)$$

After eliminating all second order terms, we are left with a closed system of asymptotic equations:

$$\left(\frac{1}{\tau} - 1\right) \left[\frac{\partial T_0}{\partial t_0} + \nabla_0 \cdot (\mathbf{u}_0 T_0) \right] = (\nabla_0^4 - 1) \nabla_0^2 T_0 - \frac{\partial}{\partial z_0} \nabla_0^2 p_0 \quad (34)$$

$$\left(\frac{\partial^2}{\partial x_0^2} + \frac{\partial^2}{\partial y_0^2} \right) p_0 = - \frac{\partial}{\partial z_0} \nabla_0^4 T_0 \quad (35)$$

$$\nabla_0^2 u_0 = \frac{\partial p_0}{\partial x_0} \quad (36)$$

$$\nabla_0^2 v_0 = \frac{\partial p_0}{\partial y_0} \quad (37)$$

The power law scaling for the heat flux is therefore $\langle wT \rangle \sim w_0 T_0 \sim \epsilon^2$. Similarly, the thermal variance and density flux are found to scale as $\langle T^2 \rangle \sim \epsilon^{3/2}$ and $\langle w\rho \rangle \sim \epsilon^{5/2}$ respectively.

In this weakly nonlinear framework, the saturation of the double-diffusive instability can be explained in terms of triad interactions between the various normal modes. Numerical simulations of the 2D equations roughly reproduce the asymptotic prediction; the scaling laws within the numerical results are estimated to be

$$\langle wT \rangle \sim \epsilon^{1.73} \quad \langle T^2 \rangle \sim \epsilon^{1.43} \quad \langle w\rho \rangle \sim \epsilon^{2.71}. \quad (38)$$

5.2 Heat–Salt fingers

For Heat–salt fingers, the Prandtl number is ≈ 7 , so we cannot neglect inertial effects. Also, the density ratio in the oceans is typically close to unity, so weakly nonlinear models are unlikely to be of much relevance. Instead, Radko [17] examined the strongly nonlinear regime, with $\delta = \sqrt{1 - R_\rho^{-1}} \ll 1$. Two dimensional numerical simulations of this regime show “modons”, small coherent vortex pairs that seem to be largely responsible for the transport of heat and salt (see Figure 10). In the limit $\delta \rightarrow 0$, an explicit solution for a

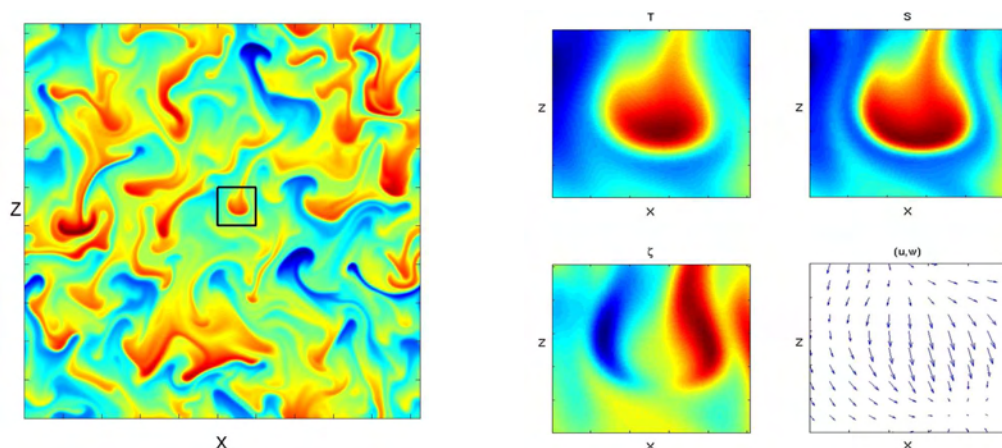


Figure 10: Strongly nonlinear heat–salt fingers in 2D, and properties of a single “modon”.

circular, rectilinearly propagating modon can be found, in polar coordinates, by expanding in powers of δ . The solutions have $w \sim \delta^{-3/4}$ and $T, S \sim \delta^{-1/4}$, so that $F_T, F_S \sim 1/\delta$, assuming that the modons dominate transport. Again, 2D numerical simulations are roughly in agreement with the asymptotic theory.

6 Thermohaline Staircases

One of the most intriguing aspects of double diffusive convection is its ability to transform smooth vertical gradients into a stepped structure consisting of mixed layers separated by thin stratified interfaces. These layers have great significance for small-scale mixing in the oceans; tracer release experiments in the North Atlantic central thermocline suggest that the presence of layers increases the effective diffusivity by salt fingers from $0.1 \text{ cm}^2 \text{ s}^{-1}$ to $1 \text{ cm}^2 \text{ s}^{-1}$ [22, 33, 27, 9]. Staircases are routinely observed in regions where the density ratio is $R_\rho < 1.9$ [21], but their origin is still poorly understood. Several questions can be raised:

- What is the origin of thermohaline staircases?
- What are the conditions for staircase formation?
- What sets the vertical scale of the layers?

6.1 Explanation of layer formation

Several hypotheses have been put forward to explain the origin of thermohaline staircases, e.g. [26, 11]. Here, we look at the suggestion by Radko [15] that layers formation arises from an instability of the flux gradient laws. A secondary instability then allows layers to merge, until the staircase achieves a stable vertical scale. This mechanism can be tested against numerical simulations of staircase formation (see Figures 11 and 12).

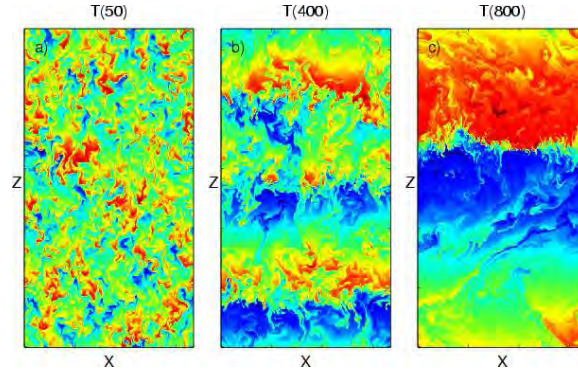


Figure 11: Formation and evolution of layers in a numerical experiment. The temperature field is shown at $t = 50$, $t = 400$ and $t = 800$.

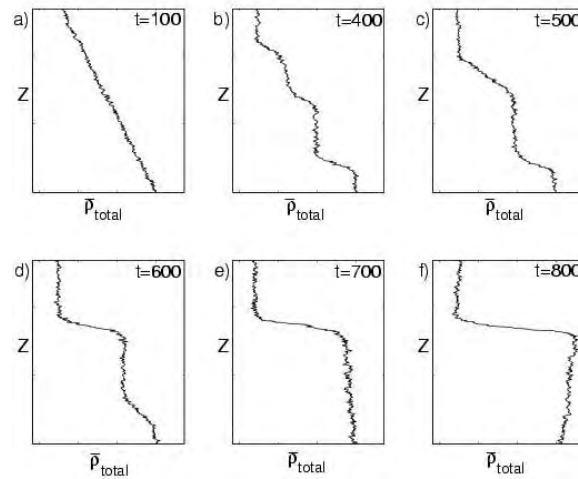


Figure 12: Total x -averaged density for the same simulation as in Figure 11 at $t = 100, 400, 500, 600, 700, 800$.

6.2 Instability of a uniform gradient

After averaging over small-scale variations of T and S , we have (1D) conservation equations

$$\frac{\partial T}{\partial t} = -\frac{\partial}{\partial z} F_T \quad (39)$$

$$\frac{\partial S}{\partial t} = -\frac{\partial}{\partial z} F_S \quad (40)$$

We suppose that the Nusselt number Nu and flux ratio γ depend only on the local density ratio R_ρ :

$$\frac{F_T}{-k_T T_z} = \text{Nu}(R_\rho) \quad (41)$$

$$\frac{\alpha F_T}{\beta F_S} = \gamma(R_\rho) \quad (42)$$

We now consider small (linear) perturbations (T_1, S_1, R_1) to a background state (T_0, S_0, R_0) with uniform gradient in T and S . The qualitative form of $\gamma(R_\rho)$ is indicated schematically in Figure 13. In the weakly unstable regime as $R_\rho \rightarrow 1/\tau$, the flux ratio must become

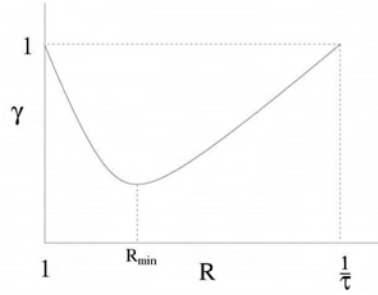


Figure 13: Dependence of the flux ratio γ on the density ratio R

unity. In the strongly unstable regime as $R_\rho \rightarrow 1$ we expect “eddy mixing” to render both temperature and salt fluxes equal. Experiments at intermediate values of R_ρ suggest that γ may obtain a minimum < 1 for some $R_\rho = R_{\min}$ [20].

A normal mode analysis for perturbations of the form $(T_1, S_1) \propto \sin(kz) \exp(\lambda t)$ yields an eigenvalue equation for the growth rate λ :

$$\lambda^2 + \lambda \left(D_2 + \text{Nu}(R_0) - D_1 \text{Nu}(R_0) R_0 - \frac{R_0 D_2}{\gamma(R_0)} \right) k^2 - D_1 \text{Nu}^2(R_0) k^4 / R_0 = 0 \quad (43)$$

where

$$D_1 = \left. \frac{\partial(1/\gamma)}{\partial R} \right|_{R_\rho=R_0} R_0 \quad (44)$$

and

$$D_2 = \left. \frac{\partial(\text{Nu})}{\partial R} \right|_{R_\rho=R_0} R_0. \quad (45)$$

So the basic uniform gradient is linearly unstable if $D_1 > 0$, i.e. if γ *decreases* with increasing R_ρ . If the dependence of γ on R_ρ is of the form shown in Figure 13 then instability is present for any $R_\rho < R_{\min}$. The physical mechanism responsible for this instability is sketched in Figure 14. With R_ρ and γ as shown, there is a convergence of heat flux in the lower domain, amplifying the perturbation to the background linear temperature gradient. In the nonlinear regime, this instability produces well-defined layers and interfaces.

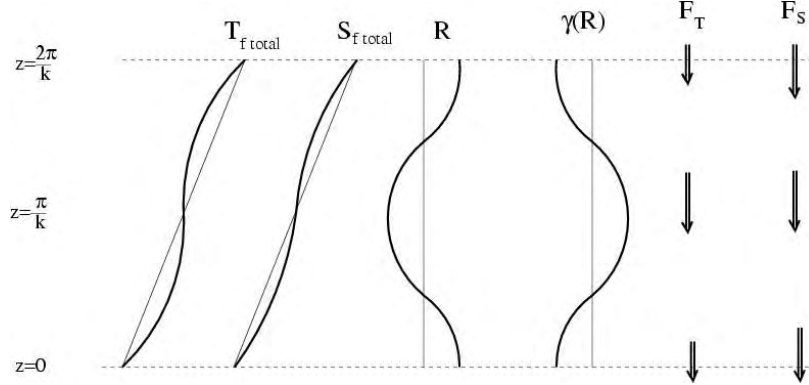


Figure 14: The physical mechanism behind the γ -instability. A decrease in γ with R results in the growth of the perturbation to a uniform T - S gradient.

To test whether variations of γ are essential for the formation of layers, we turn to the results of numerical simulations. If we define ρ_f as the horizontal average of the full (nonlinear) perturbation of density to the uniform background gradient, then we can write

$$\int \rho_f \frac{\partial \rho_f}{\partial t} dz = \int \rho_f \frac{\partial}{\partial z} (F_T - F_S) dz \quad (46)$$

$$\Rightarrow \frac{d}{dt} \int \frac{\rho_f^2}{2} dz = \int \rho_f \left(1 - \frac{F_S}{F_T}\right) \frac{\partial}{\partial z} F_T dz + \int \rho_f F_T \frac{\partial}{\partial z} \left(-\frac{F_S}{F_T}\right) dz \quad (47)$$

$$\equiv I_{\text{Nu}} + I_\gamma \quad (48)$$

where the integrals I_{Nu} and I_γ respectively quantify the effects of variation in heat flux and flux ratio. The values of I_{Nu} and I_γ were continuously recorded in the course of a numerical simulation and the results are shown in Figure 15. We see that I_{Nu} typically makes a negative contribution to the density variance, so a positive contribution from I_γ is essential for the formation of layers.

6.3 Merger of thin layers

Numerical simulations suggest that the layers formed by the γ -instability are not steady. The second process in the creation of staircases is the merger of thin layers, increasing the T and S variations across each interface. To explain the observed layer interaction, we

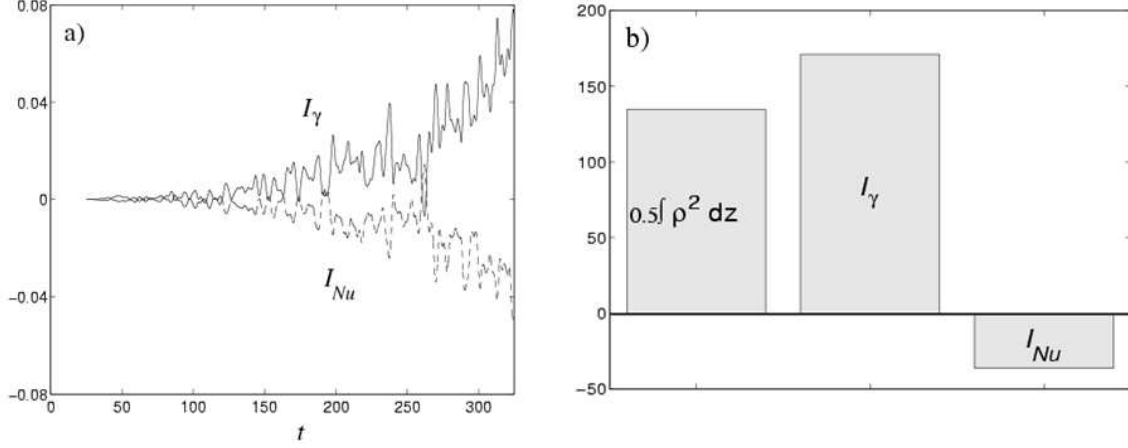


Figure 15: Balance of terms in the density variance equation (48). (a) The time evolution of I_γ and I_{Nu} , indicated by solid and dashed lines. (b) Time integrals of I_γ and I_{Nu} over the period of layer formation.

now consider the stability of a series of salt finger interfaces. As illustrated by Figure 16, we perturb the temperature profile, making alternate interfaces weaker and stronger. The

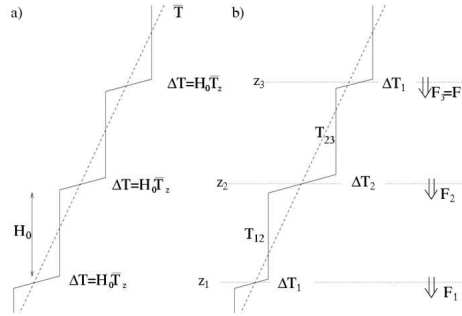


Figure 16: Temperature profiles of an infinite series of interfaces, taken from [16]. (a) Unperturbed state with equal steps (b) Perturbed state in which the T - S jumps at the even interfaces are decreased, and the jumps at odd interfaces are increased.

temperatures of the layers then evolve according to

$$H \frac{\partial}{\partial t} T_{12} = F_{T2} - F_{T1} \quad (49)$$

$$H \frac{\partial}{\partial t} T_{23} = F_{T3} - F_{T2} \quad (50)$$

We calculate the temperature fluxes across each interface using Turner's 3/4 flux law: $F_T = C(R_\rho)(\Delta T)^{4/3}$ (see §4.1). Assuming that $(\Delta T_1 - \Delta T_2) \ll \Delta T$ and $(\Delta S_1 - \Delta S_2) \ll \Delta S$

we perform a normal mode analysis, which results in an eigenvalue problem for the growth rate λ :

$$\lambda^2 + \lambda(\dots) - \frac{\partial(\gamma^{-1})}{\partial R_\rho} \bigg|_{R_\rho=R_0} \frac{64T_{0z}^{2/3}}{3H^{4/3}} C^2(R_0)k^4 = 0 \quad (51)$$

As before, there exists a positive root of this equation if $\frac{d(1/\gamma)}{dR_\rho} > 0$, so the equally stepped state is unstable if γ is a decreasing function of R_ρ .

6.4 Equilibration of thermohaline staircases

We have yet to explain the mechanism that halts the merging of the layers, and thus sets the scale of the observed steady staircases. To do this, we augment the analysis of §6.3 with convective dynamics, parameterised by the convective flux law $\text{Nu}_L = C_L \text{Ra}_\rho^{0.2}$ [7]. This situation is illustrated by Figure 17. We now find that layers become stable above a certain

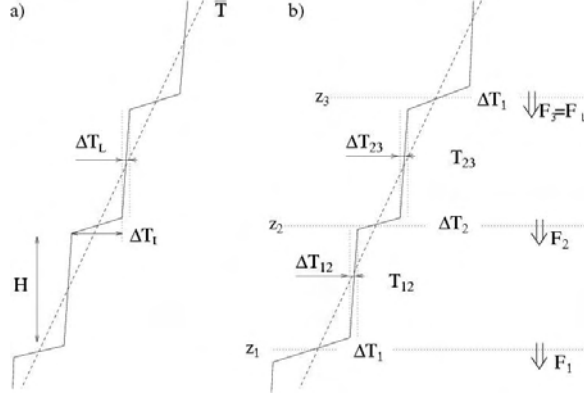


Figure 17: Temperature profiles of an infinite series of interfaces [16], including convective fluxes across each layer (c.f. Figure 16).

critical height, H_{cr} , given by

$$H_{\text{cr}} = \left(\frac{C_L}{C} \right)^{15/8} \frac{\left(\frac{R_{\text{min}}}{R_0} - 1 \right)^{9/4} \left(\frac{R_{\text{min}}}{\gamma_{\text{min}}} - 1 \right)^{1/4}}{\left(\frac{R_{\text{min}}}{\gamma_{\text{min}}} - \frac{R_{\text{min}}}{R_0} \right)^{5/2}} \left(\frac{1}{\gamma_{\text{min}}} - 1 \right)^{3/8} \left(\frac{k_T \text{Nu}}{g \alpha T_{0z}} \right)^{1/4}. \quad (52)$$

Numerical simulations of the parameterised (1D) equations [16] demonstrate both the initial instability of the uniform gradient and the secondary merging instability (see Figure 18). The merging events halt when the layers become larger than the predicted critical height H_{cr} .

7 Future challenges

Although progress has been made in quantifying the effect of double diffusion on temperature flux, we are still lacking a unified theory for the whole parameter space. Little is known

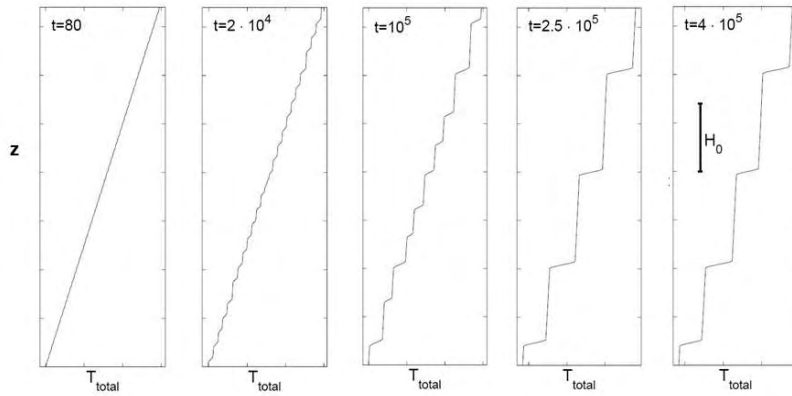


Figure 18: Results of a 1D simulation including parameterisations for both salt fingers and convection (taken from [16]).

regarding the dependence of the Nusselt number on the Lewis number or Prandtl number. The large-scale consequences of double diffusion have also not been fully studied, such as its role in controlling the T - S relation, diapycnal velocity and lateral mixing. Further analysis is required to predict and explain the magnitude and dynamics of thermohaline intrusions. Finally, there are unanswered questions regarding the modification of double-diffusive instability in the presence of shear, waves and turbulence, and the nature of the salt finger anisotropy.

References

- [1] V. W. EKMAN, *On dead water*, Scientific Results of the Norwegian North Polar Expedition, 1893-1896, 5 (1906), pp. 1-152.
- [2] R. W. GRIFFITHS AND B. RUDDICK, *Accurate fluxes across a salt-sugar finger interface deduced from direct density measurements*, J. Fluid Mech., 99 (1980), pp. 85-95.
- [3] H. E. HUPPERT AND J. S. TURNER, *Double-diffusive convection*, J. Fluid Mech., 106 (1981), p. 299.
- [4] W. S. JEVONS, *On clouds; their various forms, and protruding causes*, Sydney Magazine of Science and Art, 1 (1858), pp. 163-176.
- [5] T. M. JOYCE, *Marginally unstable salt fingers: limits to growth*, Journal of Marine Research, 40 (Suppl.) (1982), pp. 291-306.
- [6] D. E. KELLEY, *Fluxes through diffusive staircases: A new formulation*, Journal of Geophysical Research, 95 (1990), pp. 3365-3371.
- [7] R. KRISHNAMURTI, *Double diffusive transports in laboratory thermohaline staircases*, J. Fluid Mech., 483 (2003), pp. 287-314.

- [8] R. B. LAMBERT AND J. DEMENKOW, *On the vertical transport due to fingers in double diffusive convection*, J. Fluid Mech., 54 (1972), pp. 627–640.
- [9] L. S. LAURENT AND R. W. SCHMITT, *The contribution of salt fingers to vertical mixing in the north atlantic tracer release experiment*, Journal of Physical Oceanography, 29 (1999), pp. 1404–1424.
- [10] T. J. McDOUGALL AND J. R. TAYLOR, *Flux measurements across a finger interface at low values of the stability ratio*, Journal of Marine Research, 42 (1984), pp. 1–14.
- [11] W. J. MERRYFIELD, *Origin of thermohaline staircases*, J. Phys. Oceanogr., 30 (2000), pp. 1046–1068.
- [12] I. PRIKASKY, *Direct numerical simulations of the diffusive convection and assessment of its impact on arctic climate change*, Thesis, (2007).
- [13] M. R. E. PROCTOR, *Steady subcritical thermohaline convection*, J. Fluid Mech., 105 (1981), pp. 507–521.
- [14] M. R. E. PROCTOR AND J. Y. HOLYER, *Planform selection in salt fingers*, J. Fluid Mech., 168 (1986), pp. 241–253.
- [15] T. RADKO, *A mechanism for layer formation in a double diffusive fluid*, J. Fluid Mech., 497 (2003), pp. 365–380.
- [16] T. RADKO, *What determined the thickness of layers in a thermohaline staircase?*, J. Fluid Mech., 523 (2005), pp. 79–98.
- [17] ———, *The double-diffusive modon*, J. Fluid Mech., 609 (2008), pp. 59–85.
- [18] T. RADKO AND M. STERN, *Finite-amplitude salt fingers in a vertically bounded layer*, J. Fluid Mech., 425 (2000), pp. 133–165.
- [19] T. RADKO AND M. E. STERN, *Salt fingers in three dimensions*, J. Mar. Res., 57 (1999), pp. 471–502.
- [20] R. W. SCHMITT, *Flux measurements on salt fingers at an interface*, Journal of Marine Research, 37 (1979), pp. 419–436.
- [21] ———, *Form of the temperature salinity relationship in the central water: evidence for double-diffusive mixing*, Journal of Physical Oceanography, 11 (1981), pp. 1015–1026.
- [22] R. W. SCHMITT, J. R. LEDWELL, E. T. MONTGOMERY, K. L. POLZIN, AND J. M. TOOLE, *Enhanced diapycnal mixing by salt fingers in the thermocline of the tropical atlantic*, Science, 308 (5722) (2005), pp. 685–688.
- [23] C. Y. SHEN, *Heatsalt finger fluxes across a density interface*, Physics of Fluids A, 5 (1993), pp. 2633–2643.
- [24] M. E. STERN, *The “salt-fountain” and thermohaline convection*, Tellus, 12 (1960), pp. 172–175.

- [25] ———, *Lateral mixing of water masses*, Deep-Sea Research, 14 (1967), pp. 747–753.
- [26] ———, *Collective instability of salt fingers*, J. Fluid Mech., 35 (1969), pp. 209–218.
- [27] M. E. STERN, T. RADKO, AND J. SIMEONOV, *3-d salt fingers in an unbounded thermocline with application to the central ocean*, Journal of Marine Research, 59 (2001), pp. 355–390.
- [28] H. STOMMEL, A. B. AARONS, AND D. BLANCHARD, *An oceanographic curiosity: the perpetual salt fountain*, Deep-Sea Res., 3 (1956), pp. 152–153.
- [29] K. TSUBAKI, S. MARUYAMA, A. KOMIYA, AND H. MITSUGASHIRA, *Continuous measurement of an artificial upwelling of deep seawater induced by the perpetual salt fountain*, Deep-Sea Res., Part I, 54 (2007), pp. 75–84.
- [30] J. S. TURNER, *Salt fingers across a density interface*, Deep-Sea Res., 14 (1967), pp. 499–611.
- [31] G. VERONIS, *On finite amplitude instability in thermohaline convection*, Journal of Marine Research, 23 (1965), pp. 1–17.
- [32] ———, *Effect of a stabilizing gradient of solute on thermal convection*, J. Fluid Mech., 34 (1968), pp. 315–336.
- [33] ———, *Updated estimate of double diffusive fluxes in the c-salt region*, Deep Sea Research Part I: Oceanographic Research Papers, 54 (2007), pp. 831–833.
- [34] A. WILSON, *Structure and dynamics of the thermohaline staircases in the beaufort gyre*, Naval Postgraduate School Thesis, (2007).
- [35] Y. YOU, *A global ocean climatological atlas of the turner angle: Implications for double-diffusion and water-mass structure*, Deep-Sea Res., 49 (2002), pp. 2075–2093.

FELLOWS REPORTS

Models of Volcanic Tremor and Singing Icebergs

Sylvain Barbot

January 12, 2009

Abstract

We explore the possibility of flow instability through an elastic channel to explain seismic observations of tremors in such varied environments as volcanoes and icebergs. We consider the flow to be constricted in a narrow conduit and subjected to perturbations due to the propagation of compressional waves in the larger conduit. The pressure wave in the fluid are generated at the outlet of the channel by the oscillation of the walls. The growth of these self-excited perturbations may be a source of nonlinear oscillations of the wall in the channel. We explore the linear stability of various models driven by pressure gradients in the channel. We find that no linear instability persists without inertial effects in the channel. For finite yet small Reynolds number, we find growing oscillatory modes corresponding to pressure variations in the fluid flowing through the channel. Without inertial effects, the decaying oscillations in the channel have frequencies close to the natural notes of the conduit, corresponding to sound waves bouncing back and forth in the conduit. With inertial effects, however, the apparent frequencies of growing oscillatory modes are smaller than the corresponding harmonic frequencies of the conduit. The smaller apparent frequency of growing modes compared to the normal modes of a magma chamber (or reservoir, for water systems) relaxes the constraints on the linear dimension of the inferred finite bodies that host the tremor source. The proposed self-excited model may be a relevant candidate to explain observations of tremors in various geologic contexts, including at deep and shallow depths in volcanic plumbing systems and in complex water circulation in icebergs. Inertial effects in fluid flow may be responsible for the observed frequency drift which is a characteristic of tremors.

1 Introduction

Volcanoes and, as recently observed icebergs [15], are generating long period seismic activity that can persist for several minutes up to a few months. The so-called harmonic tremors are characterized by their low frequency content with a peaked spectra between 0.15 and 10 Hz. Within volcanoes, where they are generally observed, tremors often occur in connection with eruptions [2, 3] and their source originates from the shallow depth to about 50 km.

Volcanic tremors have attracted considerable attention because of their potential to constrain physical processes occurring inside volcanoes plumbing system. Tremors have been observed in numerous volcanoes including Mount Kilauea in Hawaii [16, 2], Mt. Etna [9] and Mt. St. Helens [10]. For a recent review of the observational evidence, we refer to [14]. The driving mechanism behind tremor activity is believed to involve a complex interaction between magmatic flow and surrounding bedrock in the volcano pipe system [1, 11, 3] as opposed to the brittle failure of rock that characterizes tectonic earthquakes. Tremors appear to be primarily composed of P waves [2].

Early theoretical models of tremors considered the vibration of compressible fluid-filled cracks in a layered elastic crust [7]. Aki et al. [1, 2] modeled the source of tremors to be the vibration of a crack filled with magma driven by the excess magmatic pressure. Such a jerky extension of the crack sets up a vibration, with a predominant period proportional to the linear dimension of the crack, and the amplitude proportional to the excess pressure and to the area of extension. Chouet [8] explored this assumption further with a model consisting of three elements, namely a triggering mechanism -an explosive point-source overpressure-, a resonator, and a radiator. The low frequency of tremors can be explained by normal mode oscillation in a volcanic magma chamber. For instance, the extremely long-period volcanic tremor, with periods up to 7s, observed at Mount Aso may be generated by a fluid-filled crack of modest size, a magma body 0.5 m thick and 0.5 km long [11]. The large linear dimension of required magma body to explain low frequency signals might be found in certain volcanic contexts, but may not be found systematically at shallow depths. Julian [13] showed that certain aspects of volcanic tremors such as periodic and chaotic oscillations, correlation between tremor activity and surface eruption, changes in amplitude, or frequency drifting can be explained by the flow of an incompressible viscous fluid through a channel with movable elastic walls, noting further that the obtained nonlinear process finds analogies with the excitation mechanism of musical wind instruments [4, 5]. Balmforth et al. [6] explored the stability of an incompressible flow through elastic conduits buried in a Hookean solid. They found that a critical Reynolds number is required for growing instability, and that Instabilities analogous to roll waves occur in this system.

In this study, we look for a mechanism for self-excitation of oscillations. We start with the model of Balmforth et al. [6] consisting of fluid flow through an elastically deformable channel but we couple the flow to sound waves in a larger conduit. The compressional waves in the conduit modulate the pressure at the outlet of the channel. We explore the possibility of such system to generate growing instabilities. The advantage of the proposed mechanism compared to previous models is that no outside perturbation, such as jerky over-pressure, a nearby earthquake or sudden crack opening is required to start the oscillations. In a first section, we develop the equations relevant to incompressible fluid flow in a narrow conduit with deformable walls. We obtain equations similar to lubrication theory but with inertia terms conserved. We consider the linear stability of simple models when inertial effects are ignored. In a next section, we consider the linear stability of dynamic model and look for a critical Reynolds number for onset of growing oscillations. In order to simplify an analytic treatment of the stability analysis, we consider the case of an open-funnel geometry. In an appendix, we derive the governing equations for the case of a compressible fluid.

2 Model and Governing Equations

We consider the two-dimensional flow of a Newtonian fluid in conduit of width W being constricted into a channel of height H and length L (see Fig. 1) such as $H \ll L$ and $H \ll W$. In the channel, where $0 < x < 1$, the fluid velocity $\mathbf{v} = (u \hat{\mathbf{e}}_x + v \hat{\mathbf{e}}_y)$ satisfies the Navier-Stokes equation

$$\mathbf{v}_t + \mathbf{v} \cdot \nabla \mathbf{v} = -\frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{v} \quad (1)$$

subjected to the equation of mass conservation for incompressible material

$$\nabla \cdot \mathbf{v} = 0 \quad (2)$$

and where p is the fluid pressure, ρ is the fluid density and ν is the kinematic viscosity. We consider the boundary conditions

$$\begin{aligned} u(x, h, t) &= 0 \\ v(x, 0, t) &= u_y(x, 0, t) = 0 \\ h(0, t) &= H \end{aligned} \quad (3)$$

where $h(x, t)$ is the unknown height of the channel walls. The second condition in eq. (3) is due to the symmetry about the x -axis. In the larger conduit, $L < x < L + l$, we assume that the flow is steady and only perturbed by a pressure wave, satisfying

$$p_{tt} - c^2 \nabla^2 p = 0 \quad (4)$$

where c is the sound speed in the fluid. The higher fluid velocity in the channel generates pressure fluctuations at the exit of the constricted conduit. The pressure variations are then propagated as elastic waves in the larger conduit. The wave bounces back to the channel and perturbs the mean flow, generating potential instabilities. Let us define the dimensionless parameters

$$\begin{aligned} \tilde{x} &= L x, & \tilde{y} &= H y, \\ \tilde{u} &= U u, & \tilde{v} &= \frac{U H}{L} v, \\ \tilde{p} &= \rho \nu \frac{U L}{H^2} p, & \tilde{t} &= \frac{L}{U} t \\ \tilde{c} &= C & \tilde{l} &= L l \\ \tilde{\rho} &= \rho_0 \rho & \tilde{W} &= H W \end{aligned} \quad (5)$$

where the (soon-abandoned) tilde-decorated variables have physical dimensions. Using the Reynolds number

$$R = \frac{H^2 U}{L \nu} \quad (6)$$

and the dimensionless relative velocity (sound wave speed compared to flow speed) for the conduit

$$\alpha = \frac{C}{U} \quad (7)$$

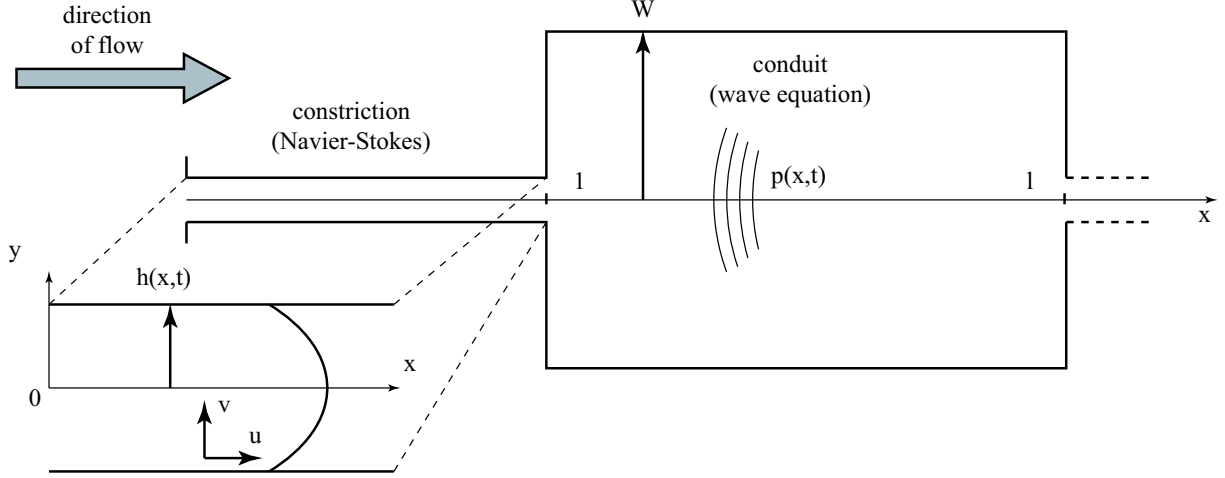


Figure 1: Geometry of the constricted flow. Fast fluid in the constricted channel generates a pressure wave in the conduit. The pressure wave bounces back to the channel and perturbs the mean flow, generating potential instability.

The dimensionless governing equations in the conduit are the wave equation

$$p_{tt} - \alpha^2 p_{xx} = 0 \quad (8)$$

and the equation of mass conservation

$$R u_t + p_x = 0 \quad (9)$$

Assuming $H/L \ll 1$, the two-dimensional Navier-Stokes equation in the channel simplifies to

$$\begin{aligned} R(u_t + u u_x + v u_y) &= -p_x + u_{yy} \\ p_y &= 0 \end{aligned} \quad (10)$$

so one writes $p = p(x, t)$. The boundary conditions are now

$$\begin{aligned} u(x, h, t) &= v(x, 0, t) = u_y(x, 0, t) = 0 \\ h(0, t) &= 1 \end{aligned} \quad (11)$$

For steady-state condition, one simply have $u_{yy} - p_x = 0$. Integrating twice and using boundary conditions of eq. (11), one obtains

$$u(\mathbf{x}, t) = \frac{1}{2} p_x (y^2 - h^2) \quad (12)$$

The average across-slot velocity

$$\bar{u}(x, t) = \frac{1}{h} \int_0^h u \, dy \quad (13)$$

is given by integration of eq. (12)

$$\begin{aligned}\bar{u}(x, t) &= \frac{p_x}{2h} \int_0^h (y^2 - h^2) dy \\ &= \frac{p_x}{2h} \left[\frac{1}{3} y^3 - h^2 y \right]_0^h = -\frac{p_x}{3} h^2\end{aligned}\tag{14}$$

The mean horizontal velocity is directly related to the pressure gradient and the wall height.

Integrating the divergence-free condition on the velocity field, one finds

$$\begin{aligned}\int_0^h \nabla \cdot \mathbf{v} &= \int_0^h u_x + v_y dy = \int_0^h u_x dy + \int_0^h v_y dy \\ &= \int_0^h u_x dy + [v]_0^h\end{aligned}\tag{15}$$

Defining $v(h) = h_t$, one gets

$$h_t + \int_0^h u_x dy = \frac{\partial}{\partial x} \int_0^h u dy - u(h) h_x + h_t = 0\tag{16}$$

but the no-shear boundary condition $u(h) = 0$ (eq. 11) leads to

$$h_t + \frac{\partial}{\partial x} \int_0^h u dy = 0\tag{17}$$

or, using the mean velocity formulation of eq. (13), one gets the expression for the conservation of mass in terms of mean horizontal flow \bar{u} and height h of the channel walls

$$h_t + (h\bar{u})_x = 0\tag{18}$$

or in terms of pressure using eq. (14)

$$h_t = \frac{1}{3} (h^3 p_x)_x\tag{19}$$

Assuming that a linear relationship exists between pressure and the wall height in the channel, eq. (19) corresponds to a nonlinear diffusion equation, which also appears in shallow water approximations [12].

2.1 Simple Models

For the sake of simplicity, we assume that the elastic wall height is uniform at equilibrium (no flow assumption). We consider that the elastic walls exert a force on the fluid proportional to the change of height (i.e., a mattress of springs approximation)

$$p = \Gamma(h - 1)\tag{20}$$

where $h = 1$ is considered the dimensionless height of the wall at equilibrium (uniform height approximation). Variable Γ is dimensionless is the corresponding physical quantity is

$$\tilde{\Gamma} = \rho \nu \frac{U L}{H^3} \Gamma\tag{21}$$

Using eqs. (20), (14) and (18), one obtains a nonlinear diffusion equation for the wall height

$$h_t = \frac{\Gamma}{3} (h^3 h_x)_x \quad (22)$$

We consider now the perturbation of the system by a sound wave propagating in the larger cavity from $x = 1$ to $x = l$. We adopt a one-dimensional approximation for the wave propagation. Using the dimensionless equation of state

$$p = \alpha^2 R \rho, \quad (23)$$

the conservation of mass

$$\rho_t + u_x = 0 \quad (24)$$

and the conservation of momentum

$$R u_t + p_x = 0, \quad (25)$$

one obtains the pressure wave equation

$$p_{tt} - \alpha^2 p_{xx} = 0 \quad (26)$$

subjected to the boundary condition $u(x = l, t) = 0$ or, equivalently, to $p_x(x = l, t) = 0$.

2.1.1 Dispersion Equation

We write the perturbation as

$$\begin{aligned} h &= 1 + h' e^{\lambda t + m x}, & 0 \leq x \leq 1 \\ p &= e^{\lambda t} P(x), & 1 \leq x \leq l + 1 \end{aligned} \quad (27)$$

where m and λ are complex and $h' \ll 1$. We further require that the pressure is continuous at the exit of the channel. Plugging eq. (27) into the nonlinear diffusion equation of eq. (22), and one obtains after linearization

$$\lambda = \frac{1}{3} \Gamma m^2 \quad (28)$$

The wall height can be written

$$h = 1 + h'_1 e^{+x\sqrt{3\lambda/\Gamma} + \lambda t} + h'_2 e^{-x\sqrt{3\lambda/\Gamma} + \lambda t} \quad (29)$$

but boundary condition at the origin $x = 0$ (see eq. (11)) gives $h'_1 = -h'_2$ so we write the wall height perturbation as follows

$$h = 1 + 2h' e^{\lambda t} \sinh \sqrt{\frac{3\lambda}{\Gamma}} x \quad (30)$$

In the larger cavity, the pressure perturbation must satisfy the wave equation. Using eqs. (26) and (27), and the separation of variables $p = e^{\lambda t} P(x)$, one gets

$$\lambda^2 P = \alpha^2 P_{xx} \quad (31)$$

solution for pressure perturbation is

$$p(x, t) = e^{\lambda t} \left(a \cosh \frac{\lambda}{\alpha} (l+1-x) + b \sinh \frac{\lambda}{\alpha} (l+1-x) \right) \quad (32)$$

with a and b real coefficient. The boundary condition $p_x(x = l, t) = 0$ demands $b = 0$. Writing the continuity of pressure at $x = 1$, one obtains

$$2\Gamma h' \sinh \sqrt{\frac{3\lambda}{\Gamma}} = a \cosh \frac{\lambda l}{\alpha} \quad (33)$$

Finally, we require the conservation of fluid flux \mathcal{F} between the two domains. The flux in the channel is given by

$$\begin{aligned} \dot{\mathcal{F}}^- &= (\bar{u} h)_t = - \left(\frac{p_x}{3} h^3 \right)_t = - \left(\frac{\Gamma}{3} h^3 h_x \right)_t \\ &= -\lambda \sqrt{\frac{\lambda \Gamma}{3}} 2h' e^{\lambda t} \cosh \sqrt{\frac{3\lambda}{\Gamma}} x + O(h'^2) \end{aligned} \quad (34)$$

where the last step is obtained by linearization of walls height perturbation. In the conduit, the rate of flux is

$$\begin{aligned} \dot{\mathcal{F}}^+ &= W u_t = -\frac{W}{R} p_x \\ &= \frac{\lambda W}{\alpha R} e^{\lambda t} a \sinh \frac{\lambda}{\alpha} (l+1-x) \end{aligned} \quad (35)$$

Equating (34) and (35), one obtains the flux continuity condition

$$\sqrt{\frac{\lambda \Gamma}{3}} 2h' \cosh \sqrt{\frac{3\lambda}{\Gamma}} + \frac{W}{\alpha R} a \sinh \frac{\lambda l}{\alpha} = 0 \quad (36)$$

Equation (36) together with eq. (33) leads to the dispersion equation

$$\sqrt{\frac{\lambda}{3\Gamma}} + \frac{W}{\alpha R} \tanh \frac{\lambda l}{\alpha} \tanh \sqrt{\frac{3\lambda}{\Gamma}} = 0 \quad (37)$$

We can see by simple inspection that no real positive part of λ can satisfy eq. (37). The simple formulation of the problem leads to a stable solution.

2.1.2 A Stability Theorem

Considering the conservation of momentum and conservation of mass equations, one can write

$$\begin{cases} u(u_t + \frac{1}{R} p_x) = 0 \\ \frac{1}{R\alpha^2} p(p_t + u_x) = 0 \end{cases} \quad (38)$$

After integrating eq. (38), one obtains

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \int_1^{l+1} R u^2 + \rho p dx &= -[up]_1^{l+1} \\ &= \frac{1}{\rho_1} (up)_{x=1} \end{aligned} \quad (39)$$

Conservation of flux at $x = 1$ ($\mathcal{F}^+ = Wu$ and $\mathcal{F}^- = -h^3 p_x/3$) and pressure eq. (20) provides us with

$$u = -\frac{\Gamma}{3W} h^3 h_x \quad (40)$$

or

$$up = -\frac{\Gamma^2}{3W} h^3 h_x (h - 1) \quad (41)$$

As conduit wall height obeys the nonlinear diffusion equation, one gets

$$\int_0^1 h_t (h - 1) dx = \frac{\Gamma}{3} \int_0^1 (h^3 h_x)_x (h - 1) dx \quad (42)$$

or

$$\frac{1}{2} \frac{d}{dt} \int_0^1 (h - 1)^2 dx = \frac{\Gamma}{3} [h^3 h_x (h - 1)]_0^1 - \frac{\Gamma}{3} \int_0^1 h_x^2 h^3 dx \quad (43)$$

Collecting terms, and using $\rho_1 = 1$, one obtains

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \left\{ \frac{\Gamma}{W} \int_0^1 (h - 1)^2 dx + \int_1^{l+1} R u^2 + \rho p dx \right\} \\ = -\frac{2\Gamma^2}{15W} \int_0^1 (h^{5/2})_x^2 dx \end{aligned} \quad (44)$$

As the right-hand-side term is negative definite, the rate of change of the total energy is negative. There is only dissipation in the system and there are no growing terms possible after perturbation.

2.2 Advection From a Pressure Gradient

We abandon the assumptions made in section 2.1 and consider in this section that fluid flow is driven by a pressure gradient γ in the channel. (We note that this model corresponds to the case of an open-funnel geometry, as developed in a later section, with the limit of neglected inertia terms.) The pressure is therefore given by

$$p(x, t) = \Gamma(h - 1) + \gamma x \quad (45)$$

where $h = 1$ is the wall height at equilibrium. The equation of continuity in the channel now reads

$$h_t = \frac{\Gamma}{3} \left(h^3 \left(h_x + \frac{\gamma}{\Gamma} \right) \right)_x \quad (46)$$

the pressure gradient introducing an effective advection term. After linearization of the perturbation solution, one obtains

$$\lambda = \frac{\Gamma}{3} m \left(m + \frac{3\gamma}{\Gamma} \right) \quad (47)$$

We write m_1 and m_2 the two solutions

$$\begin{aligned} m_1 &= -\frac{3\gamma - \sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} \\ m_2 &= -\frac{3\gamma + \sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} \end{aligned} \quad (48)$$

and the walls height in the channel can be written

$$h = 1 + 2h' \exp\left(-\frac{3\gamma}{2\Gamma}x + \lambda t\right) \sinh \frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma}x \quad (49)$$

In the conduit, we write the pressure perturbation such as

$$p(x, t) = \gamma + e^{\lambda t} P(x), \quad 1 \leq x \leq l + 1 \quad (50)$$

Solving the wave equation subjected to Neumann boundary condition at $x = l$ leads to the pressure field

$$p(x, t) = \gamma + e^{\lambda t} a \cosh \frac{\lambda}{\alpha}(l+1-x), \quad 1 \leq x \leq l+1 \quad (51)$$

Writing the equation of continuity at $x = 1$, one gets

$$2\Gamma h' e^{-3\gamma/2\Gamma} \sinh \frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} = a \cosh \frac{\lambda l}{\alpha} \quad (52)$$

We require the conservation of fluid flux \mathcal{F} between the two domains. The flux rate in the channel is given by

$$\dot{\mathcal{F}}^- = (\bar{u}h)_t = -\left(\frac{p_x}{3}h^3\right)_t = -\frac{\Gamma}{3}\left(h^3\left(h_x + \frac{\gamma}{\Gamma}\right)\right)_t \quad (53)$$

To find the flux rate about $x = 1$, we note first that

$$\begin{aligned} p_x &= \Gamma h_x + \gamma \\ h_x &= 2h' e^{\lambda t - 3\gamma/2\Gamma} \left(\frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} \cosh \frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} \right. \\ &\quad \left. - \frac{3\gamma}{2\Gamma} \sinh \frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} \right) \\ h^3 &= 1 + 6h' e^{\lambda t - 3\gamma/2\Gamma} \sinh \frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} + O(h'^2) \end{aligned} \quad (54)$$

after some algebra, one obtains

$$\begin{aligned} \dot{\mathcal{F}}^- &= -\frac{2\Gamma}{3} \lambda h' \left[\frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} \cosh \frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} \right. \\ &\quad \left. + \frac{3\gamma}{2\Gamma} \sinh \frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} \right] e^{\lambda t - 3\gamma/2\Gamma} \end{aligned} \quad (55)$$

The flux rate in the conduit does not differ from eq. (35). Equating (35) and (55), one obtains the dispersion relation

$$\begin{aligned} \left(\frac{W}{\alpha R} \tanh \frac{\lambda l}{\alpha} + \frac{\gamma}{2\Gamma} \right) \tanh \frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{2\Gamma} \\ + \frac{\sqrt{9\gamma^2 + 12\lambda\Gamma}}{6\Gamma} = 0 \end{aligned} \quad (56)$$

The leftmost quantity in parenthesis must be negative in order to a real positive solution for λ to exist. Also notice that eq. (56) simplifies to eq. (37) in the case $\gamma = 0$. Fig. 2 shows the norm of dispersion relation of eq. (56) on the complex plane with $W/\rho_1 c = 1/2$, $\Gamma = 1$ and $\gamma = -2$. The black circles in Fig. 2 correspond to the natural pulsation of the sound waves in the conduit. The zeroes of the dispersion equation -corresponding to blue circles in the contour plot- are close to the frequency corresponding to waves bouncing back and forth in the conduit. The associated eigenvalues are located on the real negative axis and correspond to decaying oscillation modes.

2.2.1 Stability Condition

Integrating the governing equations in the two domains and applying conservation of flux at $x = 1$, one finds

$$\begin{aligned} \frac{d}{dt} \left\{ \frac{\Gamma}{W} \frac{1}{2} \int_0^1 (h - H + \frac{\gamma}{\Gamma})^2 dx + \frac{1}{2} \int_1^{l+1} R u^2 + \rho p dx \right\} \\ = -\frac{\Gamma}{W} \int_0^1 h^3 h_x (h_x + \frac{\gamma}{\Gamma}) dx \\ = -\frac{\Gamma}{W} \left(\frac{2}{5} \int_0^1 (h^{5/2})_x^2 dx + \frac{\gamma}{4\Gamma} [h^4]_0^1 \right) \end{aligned} \quad (57)$$

For negative pressure gradient ($\gamma < 0$), any perturbation remains stable for $h(1) < h(0)$.

3 Fluid Flow and Sound-Wave Resonator

We now adopt a more realistic formulation of the problem where the constrained pressure gradient in the channel driving the fluid flow is compatible with the elastic deformation of the channel walls. At static equilibrium, the channel closes itself to accommodate the pressure difference between inlet and outlet. By writing the relation between pressure and wall height in the channel as follows

$$p(x, t) = \Gamma h \quad (58)$$

and using eqs. (19) and (58), the wall height satisfies the nonlinear diffusion equation

$$h_t = \frac{\Gamma}{3} (h^3 h_x)_x \quad (59)$$

subjected to the boundary conditions $p(0, t) = p_0$ and $p(1, t) = p_1 \equiv p_0 + \gamma$ at equilibrium, where γ is an effective pressure gradient. The walls height solving eqs. (19) and (58) at equilibrium ($h_t = 0$) with above-mentioned boundary conditions is

$$h(x) \equiv H(x) = \frac{1}{\Gamma} ((p_1^4 - p_0^4) x + p_0^4)^{1/4} \quad (60)$$

where we defined $H(x)$ to refer to the equilibrium, non-perturbed, solution. We now perform linearization about equilibrium as follows

$$\begin{aligned} h(x, t) &= H(x) + e^{\lambda t} h'(x), & 0 \leq x \leq 1 \\ p(x, t) &= p_1 + e^{\lambda t} P(x), & 1 \leq x \leq l+1 \end{aligned} \quad (61)$$

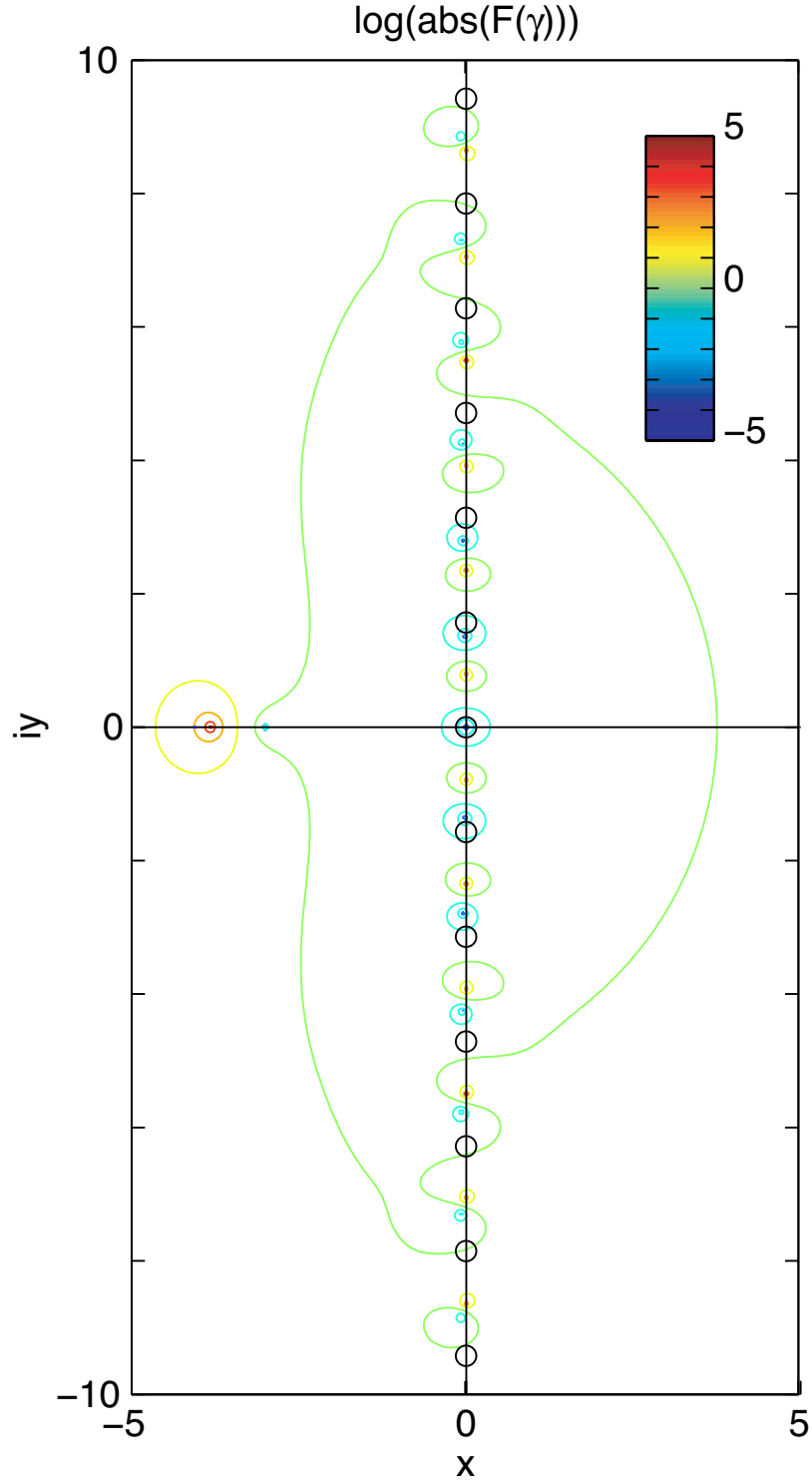


Figure 2: Contour plot on the complex plane of the dispersion relation of eq. (56) with $W/\rho_1 c = 1/2$, $\Gamma = 1$ and $\gamma = -2$. Black circles correspond to the natural notes of the conduit. There are no growing oscillatory modes.

where $h'(x)$ is small and unknown. We first note the few following results

$$\begin{aligned} h^3 &= H^3 + 3H^2 e^{\lambda t} h' + O(h'^2) \\ h^3 h_x &= H^3 H_x + (H^3 h'_x + 3H_x H^2 h') e^{\lambda t} + O(h'^2) \\ (h^3 h_x)_x &= (H^3 h'_{xx} + 6H_x H^2 h'_x - 3H_x^2 H h') e^{\lambda t} + O(h'^2) \end{aligned} \quad (62)$$

where we used the fact that H satisfies, by definition,

$$(H^3 H_x)_x = 3H_x^2 H^2 + H^3 H_{xx} = 0 \quad (63)$$

Expanding the perturbed solution in the nonlinear diffusion equation one obtains an ordinary differential for $h'(x)$,

$$\frac{3}{\Gamma} \lambda h' = H^3 h'_{xx} + 6H_x H^2 h'_x - 3H_x^2 H h' \quad (64)$$

or simply stated,

$$\mathcal{L}[h'] = \lambda h' \quad (65)$$

where the relevant boundary conditions for the eigenvalue problem are $h'(0, t) = 0$ and continuity of pressure and mass flux at $x = 1$. The flux rate at $x = 1^+$, in the larger cavity is

$$\dot{\mathcal{F}}^+ = u_t W = -\frac{1}{R} p_x W = a \lambda \frac{W}{\alpha R} e^{\lambda t} \sinh \frac{\lambda l}{c} \quad (66)$$

In the constriction, at $x = 1^-$, the flux rate is given by

$$\begin{aligned} \dot{\mathcal{F}}^- &= (\bar{u}h)_t = -\frac{1}{3} (h^3 p_x)_t = -\frac{\Gamma}{3} (h^3 h_x)_t \\ &= -\frac{\Gamma}{3} \lambda e^{\lambda t} [H^3 h'_x + 3H_x H^2 h'] + O(h'^2) \end{aligned} \quad (67)$$

Continuity of flux at $x = 1$ gives at the first order

$$a \frac{W}{\alpha R} \sinh \frac{\lambda l}{\alpha} + \frac{\Gamma}{3} (H^3 h'_x + 3H_x H^2 h') = 0 \quad (68)$$

We insist on constraining continuity of pressure at $x = 1$. We have

$$p(1^+, t) = p_1 + a e^{\lambda t} \cosh \frac{\lambda l}{c} \quad (69)$$

and

$$p(1^-, t) = \Gamma h = p_1 + \Gamma e^{\lambda t} h' \quad (70)$$

so continuity of pressure at $x = 1$ gives us

$$\Gamma h' = a \cosh \frac{\lambda l}{c} \quad (71)$$

Combining the two constraints at $x = 1$, one obtains

$$h' \frac{W}{\alpha R} \tanh \frac{\lambda l}{\alpha} + \frac{1}{3} (H^3 h'_x + 3H_x H^2 h') = 0 \quad (72)$$

Numerical inspection of solutions of the dispersion equation of eq. (64) subjected to eq. (72) at $x = 1$ reveals only solutions with negative real parts, such as the one shown in Fig. 3.

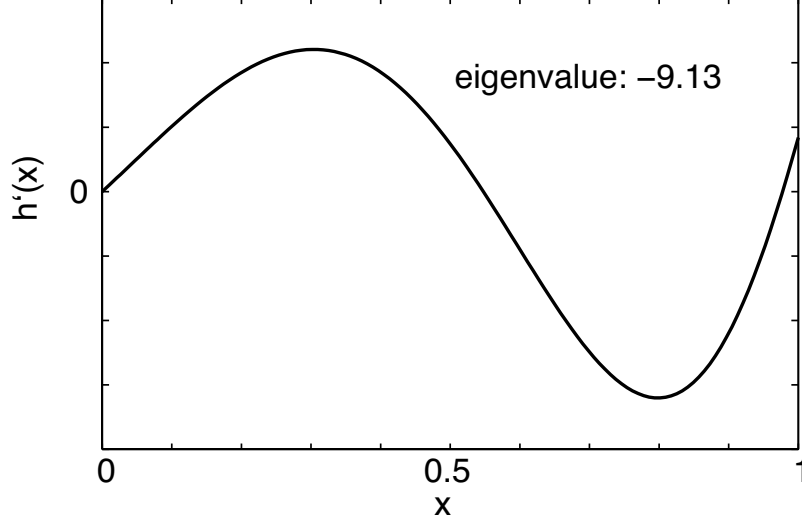


Figure 3: An eigenfunction of the dispersion function of eq. (64). Solutions correspond to decaying modes only, with negative real part eigenvalues.

3.1 Inertia Terms

We now consider the case where inertia terms in Navier-Stokes equation are small but not negligible. The governing equations adequate for lubrication theory become

$$\begin{aligned} R(u_t + u u_x + v u_y) &= -p_x + u_{yy} \\ p_y &= 0 \end{aligned} \quad (73)$$

subjected to the boundary conditions $u(x, h, t) = u_y(x, 0, t) = 0$. At leading order, when derivatives in the x -direction and acceleration can be neglected, the solution can be written

$$u(y, t) = \frac{3}{2}U \frac{h^2 - y^2}{h^2} \quad (74)$$

where $h = h(x, t)$ and by definition,

$$U = \frac{1}{h} \int_0^h u \, dy . \quad (75)$$

We assume that small inertia terms only perturb the above steady solution so that solution of eq. (73) can be of the form

$$u(x, y, t) = \frac{3}{2}U(x, t) \frac{h^2 - y^2}{h^2} \quad (76)$$

However, the eq. (76) does not provide an accurate way to evaluate u_{yy} so, when integrating eq. (73), we look for a certain projection f that allows us to avoid terms of the form u_{yy}

$$\begin{aligned} \int_0^h f(y) u_{yy} \, dy &= [f(y) u_y]_0^h - \int_0^h f'(y) u_y \, dy \\ &= [f(y) u_y]_0^h - [f'(y) u]_0^h + \int_0^h f''(y) u \, dy \end{aligned} \quad (77)$$

Setting $f''(y) = \text{cste}$ and $f(h) = f'(0) = 0$, one gets

$$\begin{aligned} f(x, y, t) &= 1 - y^2/h^2 \\ \int_0^h f u_{yy} dy &= -\frac{2}{h^2} \int_0^h u dy = -2U/h \end{aligned} \quad (78)$$

Notice that the projection kernel $f(y)$ is of the form of the steady solution of eq. (74). Projecting the lubrication approximation of the Navier-Stokes equations on $f(x, y, t)$, we look for a solution for $h(x, t)$ and $U(x, t)$. We first notice, using $u_x + v_y = 0$ and $v(0) = 0$ that one can write

$$v(x, y, t) = -\frac{3}{2}U_x y \left(1 - \frac{y^2}{3h^2}\right) - U h_x \frac{y^3}{h^3} \quad (79)$$

The projected terms are

$$\begin{aligned} \int_0^h f u_t dy &= \frac{4}{5}U_t h + \frac{2}{5}h_t U \\ \int_0^h f u_{xx} dy &= \frac{2}{5} [2U_{xx}h + 2U_x h_x + U(h_{xx} - 3h_x^2/h)] \\ \int_0^h f u_{yy} dy &= -2U/h \\ \int_0^h f u u_x dy &= \frac{36}{35}U_x U h + \frac{12}{35}U^2 h_x \\ \int_0^h f v u_y dy &= \frac{6}{35}U [3U_x h + U h_x] \\ \int_0^h -f \frac{p_x}{\rho} dy &= -\frac{2}{3}h \frac{p_x}{\rho} \end{aligned} \quad (80)$$

Defining the flux,

$$q = \int_0^h u dy = h U \quad (81)$$

the conservation of mass can be written

$$h_t + q_x = 0 \quad (82)$$

and collecting terms, one gets the coupled equations

$$\begin{cases} R \left(q_t + \frac{17}{7} \frac{q_x q}{h} - \frac{9}{7} \frac{q^2 h_x}{h^2} \right) = -\frac{5}{6} p_x h - \frac{5q}{2h^2} \\ h_t + q_x = 0 \end{cases} \quad (83)$$

Using the stress-strain relation

$$p(x, t) = \Gamma h(x, t) \quad (84)$$

one obtains the governing coupled equations

$$\begin{cases} R \left(q_t h^2 - \frac{17}{7} h_t q h - \frac{9}{7} q^2 h_x \right) = -\frac{5}{6} \Gamma h_x h^3 - \frac{5q}{2} \\ h_t + q_x = 0 \end{cases} \quad (85)$$

3.1.1 Steady-State Solution

We look for a steady solution for the wall height $H(x)$ and flux Q . Dropping time derivatives in eq. (85) and using the fact that at steady state $H_t = Q_x = 0$ and that Q is constant, one has

$$H_x \left(\frac{5}{6} \Gamma H^3 - \frac{9}{7} R Q^2 \right) + \frac{5}{2} Q = 0 \quad (86)$$

Integrating, we have

$$\frac{1}{12} \Gamma (H^4 - H_0^4) - \frac{18}{35} R Q^2 (H - H_0) + Q x = 0 \quad (87)$$

If $R \ll 1$, and the flow is constrained by the pressure p_0 at $x = 0$ and p_1 at $x = 1$, one has

$$Q = \frac{p_0^4 - p_1^4}{12 \Gamma^3} \quad (88)$$

$$H = \frac{1}{\Gamma} ((p_1^4 - p_0^4)x + p_0^4)^{1/4}$$

which correspond to solution of eq. (60), obtained for low Reynolds number. For finite Reynolds number, the steady-state flux is

$$Q = \frac{35 - \sqrt{1225 + 210 \Gamma (H_1^4 - H_0^4)(H_1 - H_0)}}{36 R (H_1 - H_0)} \quad (89)$$

3.1.2 Linear Stability Analysis

We consider perturbations of the wall height, flux and pressure of the form

$$h(x, t) = H(x) + e^{\lambda t} h'(x)$$

$$q = Q + e^{\lambda t} q'(x) \quad (90)$$

$$p = p_1 + a e^{\lambda t} \cosh \frac{\lambda}{c} (l + 1 - x)$$

where perturbations $h'(x)$ and $q'(x)$ are supposed small. Inserting eq. (90) into eq. (85) and neglecting powers of h' and q' , one obtains

$$R \left(\lambda q' H^2 - \frac{17}{7} \lambda h' H Q - \frac{9}{7} (h'_x Q^2 + 2 q' H_x) \right) \quad (91)$$

$$= -\frac{5}{6} \Gamma (H^3 h'_x + 3 H^2 H_x h') - \frac{5}{2} q'$$

The relevant boundary conditions are $h'(0, t) = q'(0, t) = 0$, continuity of pressure and flux at $x = 1$. The continuity of pressure gives

$$\Gamma h' = a \cosh \frac{\lambda l}{c} \quad (92)$$

Flux perturbation on the 1^+ side is

$$\mathcal{F}^+ = W u_t = -\frac{W}{\rho_1} p_x$$

$$= a \lambda \frac{W}{\rho_1 c} e^{\lambda t} \sinh \frac{\lambda l}{c} \quad (93)$$

On the other side, the flux perturbation is simply

$$\dot{\mathcal{F}}^- = \lambda e^{\lambda t} q' \quad (94)$$

Equating the fluxes on both sides, we get the boundary condition

$$q' - h' \frac{\Gamma W}{\rho_1 c} \tanh \frac{\lambda l}{c} = 0 \quad (95)$$

We explore the solutions of the corresponding eigenvalue problem numerically. We find a set of growing oscillatory modes associated with complex eigenvalues with positive real parts. The corresponding eigenfunctions are shown in Fig. 4.

3.2 Open-Funnel Walls

We consider now the case where the wall in the channel opens in the absence of pressure. Before fluids flow in the channel, the wall height has the equilibrium value $H^e(x)$ (see Fig. 5). And the stress-strain relation becomes

$$p(x, t) = \Gamma (h(x, t) - H^e(x)) \quad (96)$$

Using eq. (83) at equilibrium, where $h = H(x)$ and $q = Q$, we have

$$\begin{cases} R \frac{9}{7} Q^2 H_x = \frac{5}{6} \Gamma (H_x - H_x^e) H^3 + \frac{5}{2} Q \\ H_t = Q_x = 0 \end{cases} \quad (97)$$

For a general form of $H^e(x)$, the latter has no trivial solutions. We consider a special case where H_x^e is such that $H_x(x) = 0$. Under this assumption, one has $H = Q = 1$ and

$$H_x^e = \frac{3}{\Gamma} \quad (98)$$

Without inertia ($R = 0$), the governing equation (83) becomes

$$h_t = -\frac{\Gamma}{3} (h^3 (h_x - H_x^e))_x \quad (99)$$

which is similar to the formulation of section 2.2 if we write

$$\gamma = -\Gamma H_x^e \quad (100)$$

The physical grounds of the somewhat artificial pressure γ introduced in section 2.2 are the presence of opening walls in the conduit in the absence of pressure. Inspection of the dispersion relation in this case did not reveal real-positive eigenvalues so we proceed with inertia terms.

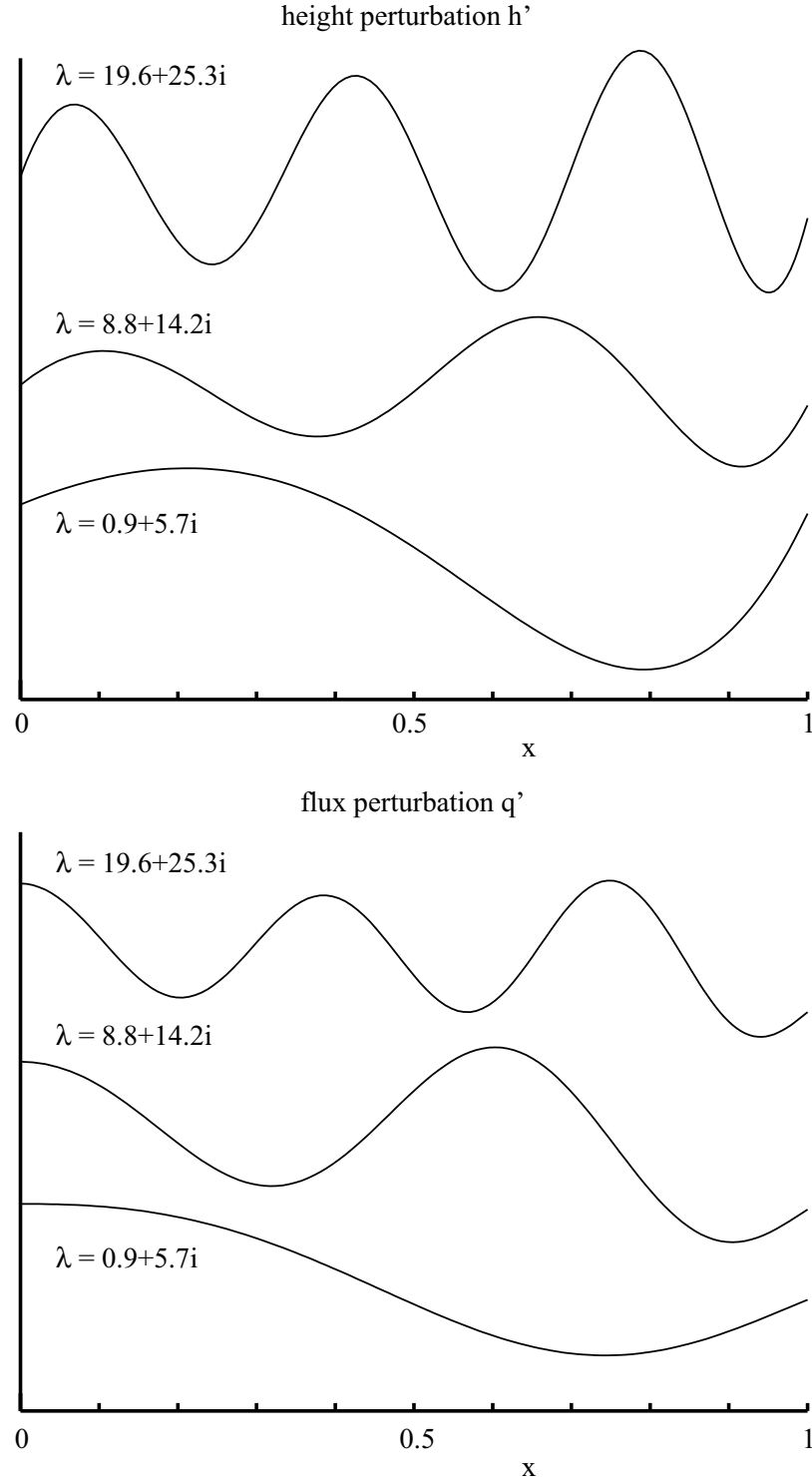


Figure 4: The first three growing eigenmodes of the system with associated eigenvalues. Top panel show the height eigenfunction $h'(x; \lambda)$ and bottom panel shows the flux eigenfunctions $q'(x; \lambda)$ in the channel.

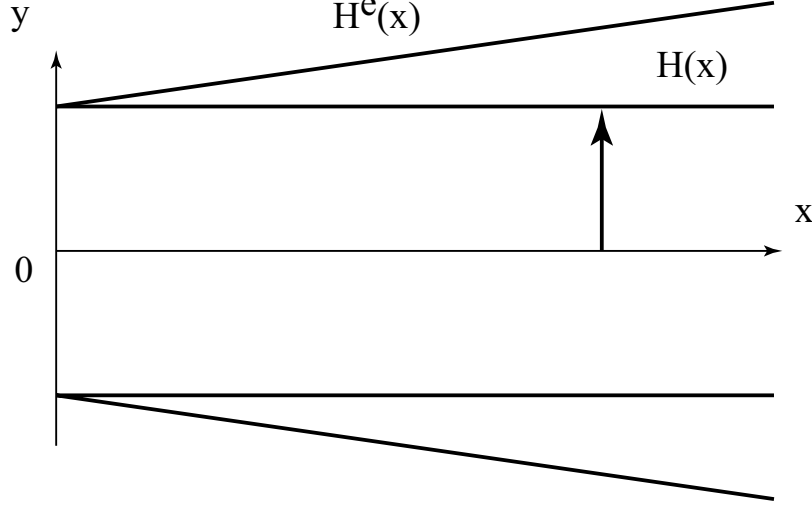


Figure 5: In the absence of fluid flow, the walls have the height $H^e(x)$. The latter is chosen such as the height of the walls is uniform when at equilibrium with steady fluid flow.

3.2.1 Stability Analysis

We write the perturbation as follows

$$\begin{aligned} h(x, t) &= 1 + h'(x) e^{\lambda t} \\ q(x, t) &= 1 + q'(x) e^{\lambda t} \\ p(x, t) &= \Gamma (1 - H^e(1)) + p' e^{\lambda t} \end{aligned} \quad (101)$$

Introducing the latter decomposition into eq. (83), one finds the system of coupled equations

$$\begin{cases} R \left(\lambda q' + \frac{17}{7} q'_x - \frac{9}{7} h'_x \right) = -\frac{5}{6} \Gamma h'_x - \frac{5}{2} q' + \frac{15}{2} h' \\ \lambda h' + q'_x = 0 \end{cases} \quad (102)$$

This result is similar to previous dispersion relations except for the newly appearing forcing term in h' on the right-hand side, so we expect to find emergent solutions, if any, for high enough Reynolds number R . The pressure continuity condition provides us with the condition

$$\Gamma h' = a \cosh \frac{\lambda l}{\alpha} \quad (103)$$

The continuity of flux at $x = 1$ leads to the second constraint

$$q' = a \frac{W}{\alpha R} \sinh \frac{\lambda l}{\alpha} \quad (104)$$

Combining the two, we obtain the boundary condition at $x = 1$

$$\lambda q' + \Gamma q'_x \frac{W}{\alpha R} \tanh \frac{\lambda l}{\alpha} = 0 \quad (105)$$

other relevant conditions are simply $h'(0) = q'_x(0) = 0$. The flux perturbation obeys the second-order homogeneous ordinary differential equation

$$\begin{aligned} \left(\frac{9}{7}R - \frac{5}{6}\Gamma\right) q'_{xx} + \left(\frac{17}{7}\lambda R + \frac{15}{2}\right) q'_x \\ + \lambda \left(\lambda R + \frac{5}{2}\right) q' = 0 \end{aligned} \quad (106)$$

The determinant of the characteristic polynomial is

$$\Delta = \left(\frac{17}{7}\lambda R + \frac{15}{2}\right)^2 - 4\lambda \left(\frac{9}{7}R - \frac{5}{6}\Gamma\right) \left(\lambda R + \frac{5}{2}\right) \quad (107)$$

Using the boundary condition $q'_x(0) = 0$, the solution flux perturbation can be written

$$q'(x) = e^{Ax} (A \sinh(Bx) - B \cosh(Bx)) \quad (108)$$

where A and B are

$$\begin{aligned} A &= -\frac{1}{2} \left(\frac{17}{7}\lambda R + \frac{15}{2}\right) / \left(\frac{9}{7}R - \frac{5}{6}\Gamma\right) \\ B &= \sqrt{\Delta} / \left(\frac{18}{7}R - \frac{10}{6}\Gamma\right) \\ C &= (A^2 - B^2) / \lambda = \left(\lambda R + \frac{5}{2}\right) / \left(\frac{9}{7}R - \frac{5}{6}\Gamma\right) \end{aligned} \quad (109)$$

Using boundary condition of eq. (105) at $x = 1$, one obtains the dispersion relation

$$\left(A + \frac{\Gamma W}{\alpha R} C \tanh \frac{\lambda l}{\alpha}\right) \tanh B - B = 0 \quad (110)$$

Fig. 6 shows a view of the dispersion equation as a function of λ on the complex plane for the parameters $\Gamma = 1$, $R = 1$, $l/\alpha = 1$ and $W/\alpha = 2$.

4 Conclusions

We investigated the potential of flow of an incompressible fluid in a channel with elastic walls coupled to a sound wave resonator to generate self-generated growing instabilities. We derived a simplified version of the Navier-Stokes equations valid for the small aspect ratio of the thin channel. We obtain equations similar to the lubrication theory, but with the inertia terms conserved. We found that simple models without flow (zero Reynolds number) did not generate instabilities. The frequency of corresponding decaying oscillation modes are close to the normal modes of sound waves in the conduit. We found linear instability corresponding to growing oscillatory modes for finite yet small Reynolds number. The small Reynolds number required for flow instability favors the occurrence of tremors in various environments, as constraints on minimum flow speed or fluid viscosity (quantities linearly related to the Reynolds number) are relaxed. The growing modes corresponding to low

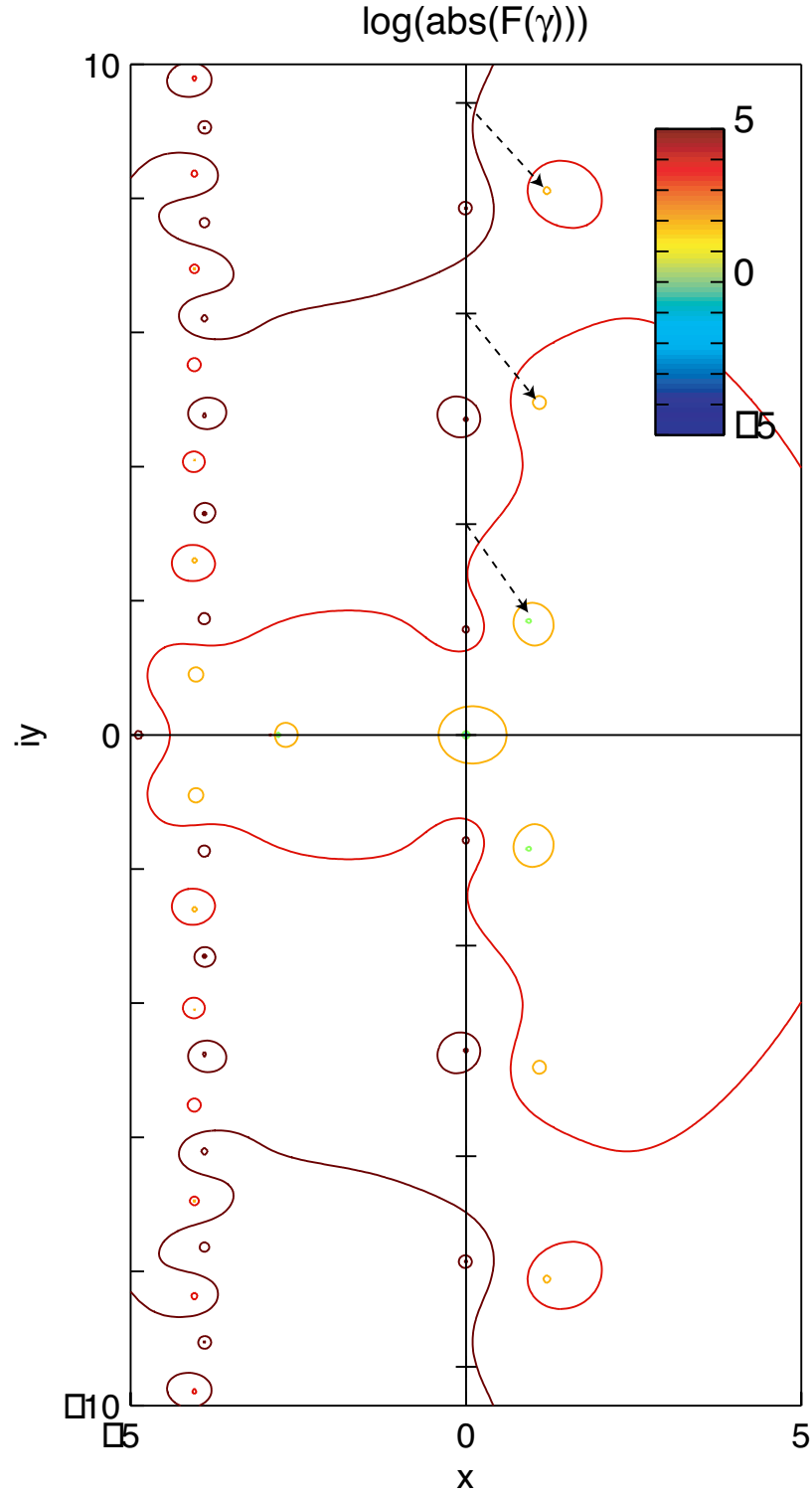


Figure 6: Map of the dispersion function as a function of eigenvalue λ on the complex plane. The zeroes indicate the position of the eigenvalues. The arrow indicate the migration of the zeroes from their position at zero Reynolds number, close to the natural notes of the conduit.

Reynolds number are associated to long time scales, giving possibly rise to long period events. The frequency of growing oscillations are smaller than the normal modes of the conduit. The smaller period of growing oscillations compare to the natural harmonics of sound wave in the conduit relaxes constraints on the linear inferred dimension of the body hosting the tremor source (water reservoir or magma chamber) and may explain the source of frequency drift observed in tremors. The proposed mechanism consisting of pressure perturbation of the fluid flow by pressure (sound) waves may be responsible for the sustained activity of volcanic tremors. The role of compressibility might affect the apparent frequency of growing modes in the channel. Implications of flow of a compressible fluid need to be investigated further as the presence of compressible fluids are relevant to the occurrence of tremors at shallow depth in volcanic plumbing systems. Our suggested mechanism may be relevant to long-period or tremor events in both magmatic and aquifer environment.

Appendix - Compressible Fluid Flow

Governing Equations

Under the assumptions adequate for lubrication theory, the Navier-Stokes equation for compressible flow reduces to

$$\begin{aligned} R(\rho u_t + \rho u u_x + \rho v u_y) &= -p_x + \eta u_{yy} \\ p_y &= \rho_y = 0 \end{aligned} \quad (111)$$

where η is the kinematic viscosity and ρ is the fluid density. The momentum equation is accompanied by the mass conservation equation

$$\rho_t + \nabla \cdot (\rho v) = 0 \quad (112)$$

where the density ρ is related to the pressure by the (non-dimensionalized) equation of state

$$p = \rho^\gamma \quad (113)$$

Approximations of the Navier-Stokes equations in section 2 are still valid for compressible flow and we have $p_y = \rho_y = 0$. Defining the volume flux

$$q = \int_0^h u dy \quad (114)$$

the conservation of mass can be written

$$(\rho h)_t + (\rho q)_x = 0 \quad (115)$$

At small Reynolds number, one has

$$q = -\frac{1}{3} p_x h^3 \quad (116)$$

and using the relationship between wall height and pressure $p = \Gamma h$, one obtains the non-linear diffusion equation

$$\left(h^{\frac{\gamma+1}{\gamma}} \right)_t = \frac{\Gamma}{3} \left(h_x h^{\frac{3\gamma+1}{\gamma}} \right)_x \quad (117)$$

At equilibrium, when $h_t = 0$, one has

$$p = [(\rho_1 p_1^4 - \rho_0 p_0^4) x + \rho_0 p_0^4]^{\frac{\gamma}{4\gamma+1}} \quad (118)$$

where H_0 , H_1 and p_0 , p_1 are the wall height and pressure respectively, at $x = 0$ and $x = 1$, respectively.

Small Reynolds Number

We evaluate the linear stability of the problem sketched in Fig. 1, with compressible flow and small Reynolds number. We write the perturbation as

$$\begin{aligned} h &= H(x) + e^{\lambda t} h'(x), & 0 \leq x \leq 1 \\ p &= p_1 + e^{\lambda t} p'(x), & 1 \leq x \leq l+1 \end{aligned} \quad (119)$$

subjected to the boundary conditions $h'(0) = p_x(l+1) = 0$ and conservation of pressure and flux at $x = 1$. The same fluid flow in the larger channel, and there the pressure p varies about the level p_1 at the exit of the contraction, so we have

$$p = p_1 + c^2 \rho \quad (120)$$

where

$$c^2 \equiv \gamma p_1^{\frac{\gamma-1}{\gamma}} \quad (121)$$

In the larger channel, the pressure obeys the wave equation subjected to $p_x(l+1) = 0$ and the solution is

$$p(x, t) = p_1 + e^{\lambda t} a \cosh \frac{\lambda}{c} (l+1-x) \quad (122)$$

Expanding the decomposition of eq. (119) into the nonlinear diffusion equation of eq. (117) and keeping the linear terms in h' , one finds

$$\frac{3}{\Gamma} \lambda \frac{\gamma+1}{\gamma} h' = \frac{3\gamma+1}{\gamma} [2H_x H^2 h'_x - H_x^2 H h'] + H^3 h'_{xx} \quad (123)$$

Notice that the latter reduces to eq. (64) for the limiting case $\gamma \rightarrow \infty$, i.e., incompressible flow. The conservation of flux at $x = 1$ is $[\rho q]_1^{1+} = 0$. The mass flux rates are

$$\dot{\mathcal{F}}^+ = (\rho_1 W u_t) = a \lambda \frac{W}{c} e^{\lambda t} \sinh \frac{\lambda l}{c} \quad (124)$$

and

$$\begin{aligned} \dot{\mathcal{F}}^- &= (\rho q)_t = -\frac{1}{3} [\rho p_x h^3]_t = -\frac{1}{3} \Gamma^{\frac{\gamma+1}{\gamma}} \left[h_x h^{\frac{3\gamma+1}{\gamma}} \right]_t \\ &= -\lambda \frac{1}{3} \Gamma^{\frac{\gamma+1}{\gamma}} \left[\frac{3\gamma+1}{\gamma} H_x H^{\frac{2\gamma+1}{\gamma}} h' + H^{\frac{3\gamma+1}{\gamma}} h'_x \right] e^{\lambda t} \end{aligned} \quad (125)$$

The conservation of pressure at $x = 1$ gives

$$\Gamma h' = a \cosh \frac{\lambda l}{c} \quad (126)$$

Using $\rho_1 = \Gamma^{\frac{1}{\gamma}} H^{\frac{1}{\gamma}}$ at $x = 1$, we obtain the boundary condition

$$h' \frac{W}{\rho_1 c} \tanh \frac{\lambda l}{c} + \frac{1}{3} \left[\frac{3\gamma + 1}{\gamma} H_x H^2 h' + H^3 h'_x \right] = 0 \quad (127)$$

Examination of numerical solutions of the eigenvalue problem does not reveal any eigenvalue with positive real parts.

Finite Inertial Terms

We consider the case where fluid acceleration is non-negligible. In the context of lubrication theory, the governing equations are

$$\begin{aligned} R(u_t + uu_x + vu_y) &= -p_x + u_{yy} \\ p_y &= \rho_y = -0 \\ \rho_t + (\rho u)_x + (\rho v)_y &= 0 \end{aligned} \quad (128)$$

We write the solution u as follows

$$u(x, y, t) = \frac{3}{2} U(x, t) \left(1 - \frac{y^2}{h^2} \right) \quad (129)$$

Integration of the conservation-of-mass equation gives

$$\rho v = -\rho_t y - \frac{3}{2} (\rho U)_x y \left(1 - \frac{y^2}{3h^2} \right) - \rho U h_x \frac{y^3}{h^3} \quad (130)$$

We now project the governing equation on $f = 1 - y^2/h^2$, in particular

$$\begin{aligned} \int_0^h f \rho v u_y dy &= \frac{6}{35} U \left[3 \rho U_x h + \rho U h_x + 3 \rho_x U h + \frac{7}{3} h \rho_t \right] \\ &= -\frac{6}{35} \frac{q}{h} \left[3(\rho h)_t + 2 \rho q \frac{h_x}{h} - \frac{7}{3} h \rho_t \right] \end{aligned} \quad (131)$$

where the last line was obtained using conservation of mass. Other terms of the governing equation are otherwise identical to those in eq. (80). After some algebra, we obtain the system of coupled equations

$$\begin{cases} R \left((\rho q)_t + \frac{17}{7} \frac{q}{h} (\rho q)_x - \frac{9}{7} \frac{q^2}{h^2} (\rho h)_x \right) = -\frac{5}{6} p_x h - \frac{5}{2} \frac{q}{h^2} \\ (\rho h)_t + (\rho q)_x = 0 \end{cases} \quad (132)$$

Using the dimension-less equation of state $p = \rho^\gamma$ and the continuity condition $p = \Gamma h$, we obtain the set of coupled equations

$$\begin{cases} R \left(\left(h^{\frac{1}{\gamma}} q \right)_t h^2 + \frac{17}{7} q h \left(h^{\frac{1}{\gamma}} q \right)_x - \frac{9}{7} q^2 \left(h^{\frac{\gamma+1}{\gamma}} \right)_x \right) \\ \hspace{15em} = -\frac{5}{6} \Gamma h_x h^3 - \frac{5}{2} q \\ \left(h^{\frac{\gamma+1}{\gamma}} \right)_t + \left(h^{\frac{1}{\gamma}} q \right)_x = 0 \end{cases} \quad (133)$$

The set of equations reduces to eq. (83) in the limit $\gamma \rightarrow \infty$.

Steady-State Solution

The expression for the wall height at steady state $H(x)$ can be obtained by noticing that $\left(H^{\frac{1}{\gamma}} q\right)_x = 0$. We obtain the following

$$\frac{1}{12}\Gamma(H^4 - H_0^4) - \frac{18}{35}Q^2R\left(H^{\frac{\gamma+1}{\gamma}} - H_0^{\frac{\gamma+1}{\gamma}}\right) - Qx = 0 \quad (134)$$

The flux Q compatible with boundary conditions on H at $x = 0$ and $x = 1$ is given by

$$Q = \frac{-35 + \left(1225 + 210R\Gamma(H_1^4 - H_0^4)\left(H_1^{\frac{\gamma+1}{\gamma}} - H_0^{\frac{\gamma+1}{\gamma}}\right)\right)^{1/2}}{36R\left(H_1^{\frac{\gamma+1}{\gamma}} - H_0^{\frac{\gamma+1}{\gamma}}\right)} \quad (135)$$

References

- [1] K. AKI, *Source mechanism of volcanic tremor: fluid-driven crack models and their application to the 1963 Kilauea eruption*, Journal of Volcanology and Geothermal Research, 2 (1977), pp. 259–287.
- [2] K. AKI AND R. KOYANAGI, *Deep volcanic tremor and magma ascent mechanism under Kilauea, Hawaii*, J. Geophys. Res., 86 (1981), pp. 7095–7109.
- [3] S. ALPARONE, D. ANDRONICO, L. LODATO, AND T. SGROI, *Relationship between tremor and volcanic activity during the southeast crater eruption on mount Etna in early 2000*, J. Geophys. Res., 108 (2003).
- [4] J. BACKUS, *Vibrations of the reed and the air column in the clarinet*, J. Acoust. Soc. Am., 33 (1961), pp. 806–809.
- [5] ———, *Small vibration theory of the clarinet*, J. Acoust. Soc. Am., 35 (1963), pp. 305–313.
- [6] ———, *Instability in flow through elastic conduits and volcanic tremor*, J. Fluid Mech., 527 (2005), pp. 353–377.
- [7] B. CHOUET, *Ground motion in the near field of a fluid-driven crack and its interpretation in the study of shallow volcanic tremor*, J. Geophys. Res., 86 (1981), pp. 5985–6016.
- [8] ———, *Excitation of a buried magmatic pipe: A seismic source model for volcanic tremor*, J. Geophys. Res., 90 (1985), pp. 1881–1893.
- [9] M. COSENTINO, S. GRESTA, G. LOMBARDO, G. PATANE, M. RIUSCETTI, R. SCHICK, AND A. VIGLIANISI, *Features of volcanic tremors on Mt. Etna (Sicily) during the marchaugust 1983 eruption*, Bull. Volcano., 47 (1984).
- [10] M. FEHLER, *Observations of volcanic tremor at Mount St. Helens volcano*, J. Geophys. Res., 88 (1983), pp. 3476–3485.

- [11] V. FERRAZZINI AND K. AKI, *Slow waves trapped in a fluid-filled infinite crack: implication for volcanic tremor*, J. Geophys. Res., 92 (1987), pp. 9215–9223.
- [12] H. E. HUPPERT AND J. E. SIMPSON, *The slumping of gravity currents*, J. Fluid Mech., 99 (1980), pp. 785–799.
- [13] B. R. JULIAN, *Volcanic tremor: Nonlinear excitation by fluid flow*, J. Geophys. Res., (1994), pp. 11859–11877.
- [14] K. I. KONSTANTINOU AND V. SCHLINDWEIN, *Nature, wavefield properties and source mechanism of volcanic tremor: a review*, J. Volcan. Geotherm. Res., 119 (2002), pp. 161–187.
- [15] C. MULLER, V. SCHLINDWEIN, A. ECKSTALLER, AND H. MILLER, *Singing icebergs*, Science, 310 (2005), p. 1299.
- [16] D. SHIMOZURU, K. KAMO, AND W. T. KINOSHITA, *Volcanic tremor of Kilauea volcano, Hawaii, during july-december, 1963*, Bull. Earth. Res. Inst., 44 (1966).

First contact in a viscous fluid

Chris Cawthorn

January 12, 2009

1 Introduction

Lubrication theory has been successfully applied to many problems in fluid dynamics. For cases where two objects slide past each other, with a thin film of fluid filling the intervening gap, it performs admirably, and its application has formed the basis for many important advancements in engineering. Furthermore, it appears in most undergraduate courses in fluid mechanics (see [1], for example). However, when used to model the coming together of two solid objects separated by a viscous fluid, the predictions of lubrication theory are decidedly unphysical.

Consider the simple example of a rigid cylinder settling under gravity towards a rigid, horizontal plane, as illustrated by Figure 1. The thickness of the gap between the cylinder and the plane is denoted by $h(x)$, where x is a horizontal coordinate measured relative to an origin situated at the intersection of the horizontal plane and the axis of symmetry. Under the approximation key to lubrication theory (the so-called ‘lubrication approximation’, which is valid provided that $\epsilon \text{Re} \ll 1$), we ignore horizontal variations relative to the much faster vertical variations when calculating the horizontal velocity, and assume that the pressure only varies in the horizontal coordinate x . We find that the horizontal velocity has a parabolic profile

$$u(x, z) = -\frac{1}{2\mu} \frac{\partial p}{\partial x} z[h(x) - z],$$

which gives rise to a horizontal volume flux

$$Q(x) = -\frac{h^3}{12\mu} \frac{\partial p}{\partial x}.$$

Balancing this against the volume that must be squeezed out due to the settling of the cylinder, we find the so-called *Reynolds equation* for the pressure gradient

$$\frac{\partial}{\partial x} \left[-\frac{h^3}{12\mu} \frac{\partial p}{\partial x} + xR\dot{\epsilon} \right] = 0 \quad (1)$$

The dominant contribution to the upward resistive force on the cylinder comes from the pressure force

$$F_z = \int p \, dx \propto \frac{\mu R \dot{\epsilon}}{\epsilon^{3/2}}.$$

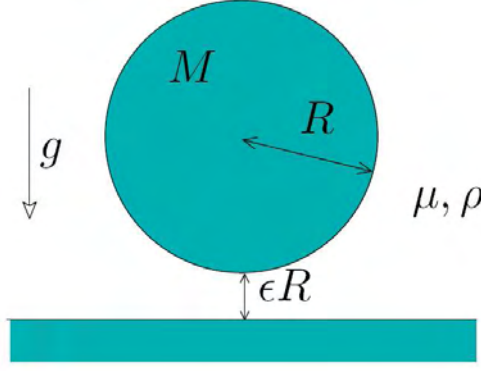


Figure 1: A simple application of lubrication theory. A rigid cylinder settles toward a rigid, horizontal plane under gravity.

The integration of the Reynolds equation (1) is straightforward, and can be well approximated analytically by extending the range of integration to $-\infty < x < \infty$, and approximating the gap thickness $h(x)$ by a parabola. The important result of this analysis is the scaling of the vertical force. Balancing this resistive force against the cylinder's weight (neglecting inertia, consistent with the lubrication approximation), we find that

$$Mg \propto \dot{\epsilon} \epsilon^{3/2} \quad (2)$$

$$\Rightarrow \epsilon(t) \propto (t - t_c)^{-2}. \quad (3)$$

So, according to lubrication theory, the cylinder (or indeed any object with a locally parabolic shape) takes an infinite length of time to make contact with the plate¹. Such a result is clearly unphysical. It is the purpose of this report to address some possibilities for resolving this inconsistency.

Many physical effects could be introduced in order to attempt to capture finite-time contact. Perhaps small asperities on the solid surfaces will provide sharper, more streamlined points of contact between the two objects. Perhaps elasticity in the solid objects or compressibility in the fluid will provide a mechanism by which the large lubrication pressures can be mitigated. If these continuum scale effects are insufficient, then perhaps an explanation in terms of long range (e.g. Van der Waals) forces can be sought. In this report, we investigate the effects of roughness, elasticity and compressibility on settling problems, and see if contact in finite time can be achieved.

2 Roughness

Real surfaces are seldom perfectly smooth, as shown by Figure 2. When considering the effect of surface roughness on settling problems, we should first point out that the lubrication approximation itself is not the sole cause of the unphysical infinite contact time. A theorem of Gérard-Varet & Hillairet [2] states that, for a sufficiently smooth object settling toward

¹For particularly flat shapes with zero local curvature, a similar analysis can be performed – see [5]

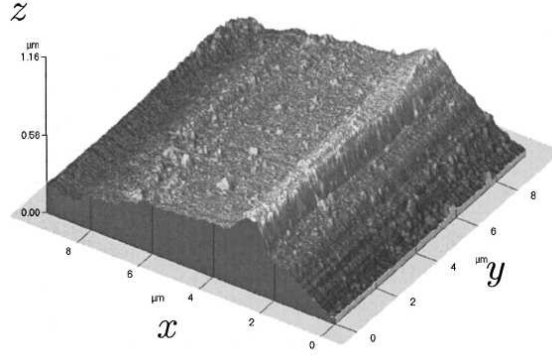


Figure 2: Atomic force micrograph of a roughness on a machined surface. Note rough asperities on the scale of microns. Picture copied from [3].

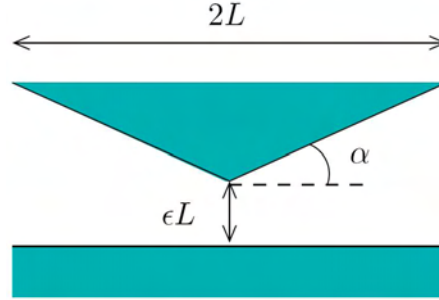


Figure 3: Diagram and notation for falling wedge

a sufficiently smooth plane under a constant body force, even the full, two-dimensional, Navier-Stokes equations require that contact will not occur in finite time. In order to resolve the problem without adding new physical processes (such as elasticity or compressibility, as discussed in §§3-4), we must insist that the settling objects contain sharp regions. By sharp, we mean regions where the first derivative of the surface function is discontinuous. In this section, we first consider the dynamics of a single sharp wedge settling in isolation, then move on to consider whether sharp roughness superimposed on a cylinder can change the settling dynamics as it rolls down an inclined plane.

2.1 Settling of a sharp wedge

In order to investigate the effect of a sharp asperity on contact, we consider the model problem of a falling wedge, as illustrated in Figure 3. The wedge is symmetric about $x = 0$, has width $2L$ and is pitched at angle α to the horizontal. We denote the closest approach of the wedge by ϵL . Assuming that the resulting flow has sufficiently small Reynolds number, we aim to solve the problem in a quasi-static manner, scaling out the settling velocity $\dot{\epsilon}$, and using the Stokes equations to find the resulting dimensionless hydrodynamical force F_Z as

a function of ϵ . We then use instantaneous force balance to equate the effective weight Mg' of the wedge with this hydrodynamic resistance. In this section, we detail two approximate analytic solutions for the Stokes velocity field, and present an exact numerical solution for comparison. Throughout this section, we shall nondimensionalise lengths with L , velocities with $\dot{\epsilon}L$, time with $\dot{\epsilon}^{-1}$, and pressure with $\mu\dot{\epsilon}$.

2.1.1 Lubrication solution

We begin by considering the flow obtained by applying the lubrication approximation to the flow beneath the falling wedge. The local thickness of the fluid layer is given by,

$$H(X) = \epsilon + |X| \tan \alpha. \quad (4)$$

Away from the sharp vertex at $X = 0$, we can expect the lubrication approximation to be reasonable provided that $\tan \alpha \ll 1$. For now, we shall not worry about the rapid horizontal variations in the vicinity of the vertex, except to note that the lubrication approximation should not be expected to perform well in this region.

After obtaining the Poiseuille velocity profile typical of lubrication flows

$$U(X, Z) = -\frac{1}{2} \frac{\partial P}{\partial X} Z [H(X) - Z], \quad (5)$$

we can integrate across the narrow gap to find the volume flux

$$Q(X) = \int_0^{H(X)} U(X, Z) dZ = -\frac{1}{12} \frac{\partial P}{\partial X} H^3$$

and use the local continuity relation to determine the Reynolds equation

$$\frac{\partial}{\partial X} \left[X - \frac{1}{12} \frac{\partial P}{\partial X} H^3 \right] = 0. \quad (6)$$

By choosing the reference pressure so that we may take $P(X = \pm 1) = 0$, we can integrate equation (6) in order to determine the pressure gradient

$$\frac{\partial P}{\partial X} = \frac{12X}{H^3}. \quad (7)$$

We may now use (7) to calculate the upward resistive force

$$F_Z = \int_{-1}^1 P dX = - \int_{-1}^1 X \frac{\partial P}{\partial X} dX = -\frac{24}{\tan^3 \alpha} \left[\log \epsilon - \frac{3}{2} + \mathcal{O}(\epsilon) \right], \quad (8)$$

and balance this against the weight of the wedge to determine the equation governing its vertical motion. Restoring dimensions, we have

$$\mu \dot{\epsilon} L \left[\log \epsilon - \frac{3}{2} + \mathcal{O}(\epsilon) \right] = -\frac{Mg' \tan^3 \alpha}{24}. \quad (9)$$

Integrating (9), we find that the minimum separation ϵ is governed by

$$\epsilon \log \epsilon \sim \frac{Mg' \tan^3 \alpha}{24\mu L} (t_1 - t), \quad (10)$$

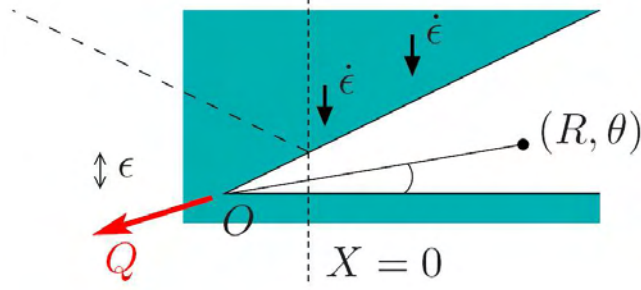


Figure 4: Diagram and notation for flow in a corner. The dashed line illustrates the outline of the wedge in Figure 3, and the symmetry axis $X = 0$.

as $\epsilon \rightarrow 0$. This suggests that the wedge will make contact with the horizontal surface at the finite time $t = t_1$ (here a constant dependent upon the initial conditions), suggesting that sharp edges might be a way to achieve contact in finite time. However, since the lubrication approximation can not be expected to hold near the vertex $X = 0$, we aim to verify this conclusion by constructing a more reasonable approximation to the Stokes equations.

2.1.2 Outer Stokes solution

We now aim to determine an approximation to the Stokes flow beneath the wedge by considering the similar problem of flow in a corner, as illustrated by Figure 4. We aim to find the Stokes flow in a corner of angle α , with velocity boundary conditions corresponding to those in the falling wedge problem. In polar coordinates (R, θ) with the origin indicated, we define a streamfunction $\psi(R, \theta)$ such that $u_R = R^{-1} \partial_\theta \psi$ and $u_\theta = -\partial_R \psi$, write the Stokes equations and boundary conditions in the form

$$\nabla^4 \psi = 0 \quad \text{in } 0 < \theta < \alpha \quad (11)$$

$$\psi(R, 0) = \frac{\partial \psi}{\partial \theta}(R, 0) = 0 \quad (12)$$

$$\frac{\partial \psi}{\partial R}(R, \alpha) = -\cos \alpha \quad (13)$$

$$\frac{\partial \psi}{\partial \theta}(R, \alpha) = R \sin \alpha. \quad (14)$$

The boundary conditions (12-14) simply represent the no-slip conditions $\mathbf{u} = 0$ on $\theta = 0$ and $\mathbf{u} = \dot{\epsilon} \hat{\mathbf{z}}$ on $\theta = \alpha$ in dimensionless variables. As stated, this biharmonic problem does not have a unique solution. If we insist that the velocity be bounded as $R \rightarrow \infty$ then we could expect to find a solution of the form

$$\psi(R, \theta) = f(\theta) + Rg(\theta). \quad (15)$$

The R -independent term in (15) will correspond to a line source or sink located at the origin. This term is not necessary when aiming to find a solution to the equations (11-14),

but has been included as a means of matching up the volume fluxes in this corner flow problem with those expected in the falling wedge problem. By symmetry, we expect that there will be no horizontal volume flux through $X = 0$ in the falling wedge. In this problem we recreate this boundary condition by selecting the strength Q of the sink at the origin so that there is no flux through $R = \epsilon/\sin \alpha$. In other words, we add the additional boundary condition

$$\psi\left(\frac{\epsilon}{\sin \alpha}, \theta\right) = 0 \quad (16)$$

to the equations (11-14). The equations can then be solved easily to yield the solutions

$$f(\theta) = \epsilon \cot \alpha [(\cos 2\alpha - 1)(\cos 2\theta - 1) + \sin 2\alpha(\sin 2\theta - 2\theta)], \quad (17)$$

$$g(\theta) = \frac{(\alpha + \sin \alpha \cos \alpha)(\theta \cos \theta - \sin \theta) + \theta \sin^2 \alpha \sin \theta}{\sin^2 \alpha - \alpha^2}. \quad (18)$$

It should be noted that this solution is *not* an exact solution for flow beneath the falling wedge discussed earlier. In particular, it does not allow for the streamfunction to be antisymmetric about the vertex of the falling wedge. However, it obeys all of the same boundary conditions away from the vertex of the wedge, so should be an outer solution, valid further than an $\mathcal{O}(\epsilon)$ distance from the vertex. It is essentially the leading-order contribution to the streamfunction in the limit $\epsilon \rightarrow 0$. Note also that this approximation is valid for arbitrary inclination α , unlike the lubrication approximation discussed in §2.1.1. Indeed, as $\alpha \rightarrow \pi/2$, this solution will exactly recreate the asymmetry needed at $X = 0$ in the falling wedge problem.

Having obtained this solution, we could consider the resistive force that the associated fluid flow should exert on the boundary ($R > \epsilon/\sin \alpha, \theta = \alpha$), corresponding to half of the wedge in Figure 3. As we expect the horizontal component of this force to vanish by symmetry, we consider only the vertical force on the whole wedge, namely

$$F_Z = 2 \int_{\epsilon/\sin \alpha}^{1/\cos \alpha + \epsilon \sin \alpha} \hat{\mathbf{z}} \cdot \boldsymbol{\sigma} \cdot \mathbf{n} dR, \quad (19)$$

$$= 2 [(g + g'') \sin \alpha + (g' + g''') \cos \alpha] \log\left(\frac{\epsilon}{\sin \alpha}\right) + 2 [f'' \sin \alpha + (f' + f''') \cos \alpha] + \mathcal{O}(\epsilon), \quad (20)$$

where the various derivatives of f and g are evaluated at $r = \alpha$. Substituting for these using the expression (17-18), we find that the vertical force is given by

$$F_Z = 2 \left(\frac{2\alpha + \sin 2\alpha}{\alpha^2 - \sin^2 \alpha} \right) \log\left(\frac{\epsilon}{\sin \alpha}\right) - 8 \cos \alpha (1 + \cos^2 \alpha) + \mathcal{O}(\epsilon). \quad (21)$$

Considering only the leading-order term, we can set this equal to the weight of the wedge and restore dimensions to find that, according to this solution

$$\epsilon \log \epsilon \sim \frac{Mg'(\sin^2 \alpha - \alpha^2)}{2\mu L(2\alpha + \sin 2\alpha)}(t_2 - t), \quad (22)$$

as $\epsilon \rightarrow 0$. Once again, the wedge will make contact with the horizontal plane in finite time, at $t = t_2$. It is very interesting to note that we obtain the same functional form as in (10),

which was obtained by lubrication theory. Furthermore, the prefactors in equations (10) and (22) have the same leading-order behaviour for $\alpha \ll 1$. This seems to be an encouraging sign that finite-time contact may be possible even with the full Stokes equations, but we must note that this outer solution does not correctly describe the flow in the vicinity of the sharp point. In order to fully demonstrate that finite-time contact *is* possible with this sharp wedge we must resort to a full, numerical solution of the Stokes equations.

2.1.3 Full Stokes solution

In this section, we obtain a numerical solution for the streamfunction $\psi(x, z)$ describing the Stokes flow beneath the falling wedge of Figure 3. For conditions on the vertical boundaries, we use the antisymmetry of ψ at $X = 0$, and match to the outer solution (15,17,18) at the edge of the wedge. We therefore solve the biharmonic problem

$$\nabla^4 \psi = 0 \quad \text{in } -1 < X < 1, 0 < Z < H(X) \quad (23)$$

$$\psi(X, 0) = \frac{\partial \psi}{\partial Z}(X, 0) = \frac{\partial \psi}{\partial X}(X, H(X)) = 0, \quad (24)$$

$$\frac{\partial \psi}{\partial X}(X, H(X)) = -1, \quad (25)$$

$$\psi(X, Z) = -\psi(-X, Z), \quad (26)$$

$$\psi(1, Z) = \psi_{\text{outer}}(1, Z), \quad (27)$$

where $\psi_{\text{outer}}(X, Z)$ is the outer solution given by (15,17,18), using the coordinate transformation

$$R = \sqrt{(X + \epsilon \cot \alpha)^2 + Z^2}, \quad \theta = \tan^{-1} \left(\frac{Z}{X + \epsilon \cot \alpha} \right). \quad (28)$$

The asymmetry constraint (26) both allows us to consider only the region $0 < X < 1$, and provides boundary conditions on the streamfunction at $X = 0$

Our numerical solution of equations (23-28) was performed by first mapping the region onto the unit square, via the transformation

$$\xi = X, \quad \zeta = \frac{Z}{H(X)}, \quad (29)$$

and transforming the equations and boundary conditions accordingly. It is then a simple matter to discretise the equations on a uniform (ξ, ζ) grid by using a second-order finite difference scheme. The numerical solution then becomes a simple linear system, which can be inverted either exactly via LU factorization, or approximately by an iterative method that takes advantage of the sparsity of the matrix defining the problem.

After testing the code by using several exact solutions of the biharmonic equation, and confirming that the error decreases as $\mathcal{O}(\Delta X^2)$, we were able to calculate the streamfunction (and hence the velocity field) for a variety of angles α . Two examples are shown in Figure 5. Observe that, for smaller angles ($\alpha = \pi/12$ in Figure 5), the horizontal velocity field exhibits the parabolic shape (5), as obtained under the lubrication approximation, throughout most of the domain. For a sharper wedge ($\alpha = \pi/3$ in Figure 5), the horizontal velocity is concentrated more towards the horizontal surface, more like the outer solution described

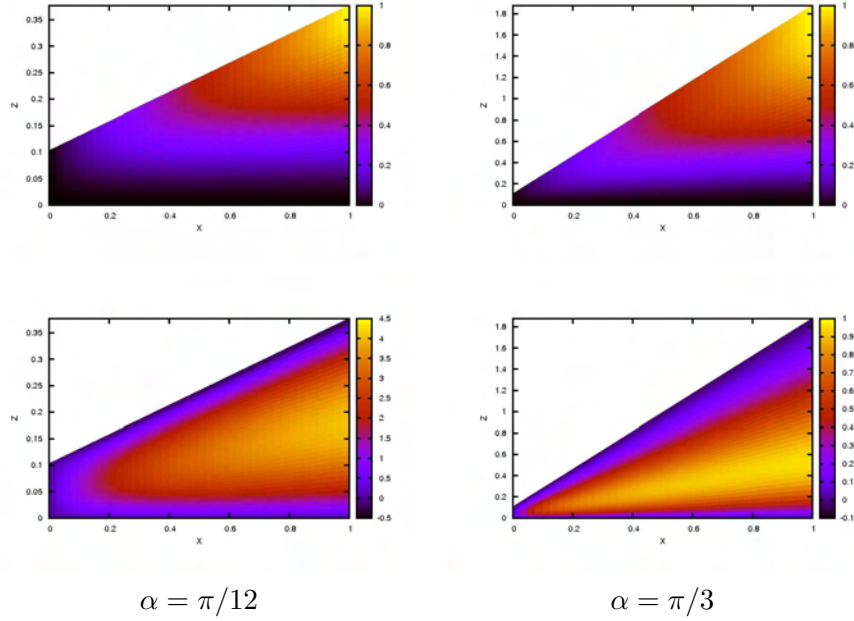


Figure 5: Plot of the streamfunction ψ and horizontal velocity $U = \partial_Z \psi$ beneath the falling wedge for $\alpha = \pi/12$ and $\alpha = \pi/3$. Note that $\epsilon = 0.1$ in both cases.

in §2.1.2. Both examples shown here indicate, therefore, that the approximate solutions considered previously can reasonably represent the main features of the flow for large and small α , respectively.

Given the streamfunction and velocity field, we can calculate the pressure and stress fields, allowing us to compute the vertical force acting on the wedge

$$F_Z = \int \hat{\mathbf{z}} \cdot \boldsymbol{\sigma} \cdot \mathbf{n} dS \quad (30)$$

$$= \int_{-L}^L \left[\cos \alpha \left(P - \frac{\partial U}{\partial X} \right) + \frac{1}{2} \sin \alpha \left(\frac{\partial U}{\partial Z} + \frac{\partial V}{\partial X} \right) \right]_{Z=H(X)} dX \quad (31)$$

$$= 2 \int_0^L \left[-\cos \alpha \left(X \frac{\partial P}{\partial X} + \frac{\partial U}{\partial X} \right) + \frac{1}{2} \sin \alpha \left(\frac{\partial U}{\partial Z} + \frac{\partial V}{\partial X} \right) \right]_{Z=H(X)} dX, \quad (32)$$

after integrating by parts, using the pressure boundary condition $P(\pm L) = 0$, and observing that all of the quantities in the integrand are even functions of X . Plots of the vertical force as a function of ϵ are shown in Figure 6. We observe that, in both the case of a relatively bluff wedge ($\alpha = \pi/12$) and a relatively sharp wedge ($\alpha = \pi/3$), the small ϵ behaviour of the force is a linear function of $\log \epsilon$, as predicted by our analytical results. Furthermore, the slope of this linear function agrees well with that predicted by the outer solution in both cases. However, in the small- α case, the lubrication solution (8) provides a better prediction of the constant term that forms the first correction. In the high- α example, the force predicted by the outer solution matches the numerical force extremely closely, but

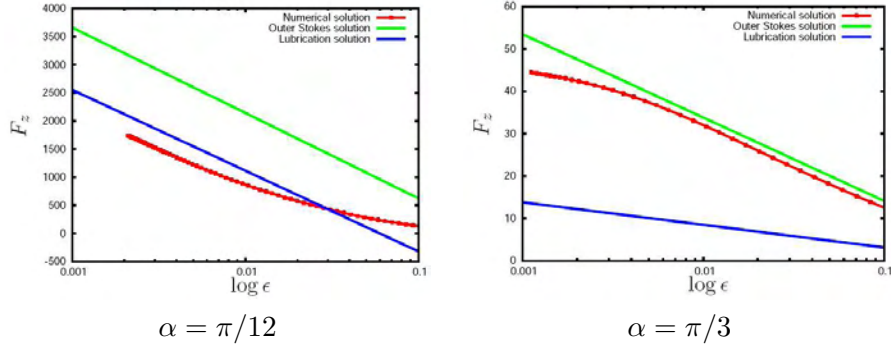


Figure 6: Plots of the vertical force F_Z , as determined by lubrication theory (8), the outer Stokes solution (21), and the numerical solution, via (32). First-order (constant) corrections have been included in the asymptotic results. Two wedge angles ($\alpha = \pi/12$ and $\alpha = \pi/3$) are shown.

it should be noted that the agreement between numerical and analytical results degrades as ϵ becomes very small. We believe that this represents a problem with the numerical resolution of our numerical scheme when the separation ϵ is very small. In this limit, the transformation (29) is near-singular in the vicinity of $X = 0$, so we should expect to need to use a higher spatial resolution in this region. This remains as potential future work.

To conclude our discussion of this problem, we make two important assertions. Firstly, both our analytical and numerical results support the hypothesis that objects with sharp asperities can make contact with a smooth surface in finite time. This suggests that roughness may indeed be an important consideration in determining settling dynamics at small lengthscales. However, one could argue that no physical surface is perfectly rough, and must be rounded on some lengthscale. Some preliminary numerical experiments with a rounded wedge were carried out, and the behaviour of the vertical force examined. For relatively large separations, the effect of rounding the corner is negligible, and the force remains proportional to $\log \epsilon$ for moderate ϵ . However, for very small lengthscales (on the order of the amount by which the corner was rounded), we observe that the force scales algebraically with ϵ . In other words, the pressure force from beneath the smoothed corner is large enough to dominate the logarithmic behaviour that arises due to the straight-edged wedge.

The second statement that we make is that, in this case at least, lubrication theory predicts the correct functional dependence of the vertical force on the minimum separation ϵ , though it is quantitatively incorrect for larger angles, as we should expect. This seems to suggest that any corrections made to the flow field due to considering the full Stokes equations are mere details that do not qualitatively affect the settling dynamics of sharp objects. For the remainder of this document, we shall use this result to justify the use of only lubrication theory in the following calculations.

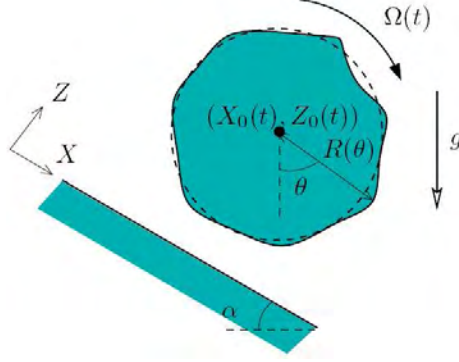


Figure 7: Diagram and (dimensionless) notation for the rough cylinder problem.

2.2 Motion of a rough cylinder

We now turn our attention to a different problem involving the effect of roughness on settling. We aim to investigate the effect of asperities on an object that is able to rotate as it settles. Should an object fall preferentially with its sharp edges pointing downwards, or will it orientate itself in such a way as to make the roughness less significant? In order to attempt to answer this question, we consider the model problem of a rough, solid cylinder settling on an inclined plane under the influence of gravity in a viscous fluid, as shown in Figure 7.

In our set-up of the problem, the cylinder has a radius

$$r(\theta) = R_0 R(\theta) = R_0 [1 + \eta(\theta)] \quad (33)$$

where θ is the polar angle measured relative to the normal to the inclined plane, as shown. The centre of the unperturbed cylinder is located at $(x_0(t), z_0(t))$, measured in Cartesian coordinates parallel and perpendicular to the plane, respectively. The rigid body motion of the cylinder can therefore be described by a translational velocity (\dot{x}_0, \dot{z}_0) and an angular velocity ω measured about this centre point. The plane itself is inclined at angle α to the horizontal, and the cylinder has effective mass $M'g$. We also define Cartesian coordinates (x, z) to be locally parallel and perpendicular to the inclined plane, with the line $x = 0$ passing through the centre of the undisturbed cylinder.

2.2.1 Governing equations

In order to describe the motion of the rough cylinder, we aim to calculate the hydrodynamic force and torque acting on the cylinder, each as a function of the rigid-body motion of the cylinder. We assume that the Reynolds number of the flow is sufficiently small to allow us to apply the Stokes equations. Linearity will then imply that we can express these functions as a linear system of the form

$$(F_X, F_Z, G)^T = \mathbf{M}(\dot{x}_0, \dot{z}_0, \omega)^T$$

where $\mathbf{F} = (F_X, F_Z)$ is the force, and G the torque exerted on the cylinder by the fluid. \mathbf{M} is a resistance matrix expressing the effect of geometry on the linear system. We aim to construct \mathbf{M} by assuming that the dominant contributions to both the force and torque come from the thin gap between the cylinder and the plane, and applying the lubrication approximation in this gap. Even for ‘sharp’ roughness, the results of §2.1 suggest that this approximation will not qualitatively change the dynamics of the cylinder when compared with a full Stokes solution under the cylinder. By then making a quasi-stationary approximation, reflecting the instantaneity of Stokes flow, we equate the hydrodynamic force and torque to those exerted by gravity, and solve for the velocity and angular velocity of the current configuration by inverting the resistance matrix \mathbf{M} . Before proceeding with the derivation of \mathbf{M} , we nondimensionalise the physical variables by writing

$$(x, z, x_0, z_0) = R_0(X, Z, X_0, Z_0), \quad (u, \dot{X}_0, \dot{Z}_0) = \frac{M'g}{\mu}(U, \dot{X}_0, \dot{Z}_0), \quad (\omega, t^{-1}) = \frac{M'g}{\mu R_0}(\Omega, T^{-1}).$$

In order to use the lubrication solution in the Cartesian coordinates (X, Z) , we must first express the thickness $H = h/R_0$ of the gap beneath the cylinder as a function of the downslope coordinate X . A simple geometrical calculation leads to the relationship

$$X = [1 + \eta(\theta - \phi)] \sin \theta, \quad (34)$$

where $\phi = \int \Omega dT$ describes the current orientation of the cylinder. The gap thickness can then be expressed as

$$H(X) = Z_0 - (1 + \eta(\theta - \phi) \cos \theta) = Z_0 - \sqrt{(1 + \eta(\theta - \phi))^2 - X^2} \quad (35)$$

Assuming that the roughness $\eta \ll 1$, we approximate the relationship (34) by taking $\theta = \sin^{-1} X$ in (35).

As is usual for lubrication flows, the horizontal velocity profile will be parabolic, and of the form

$$U(X, Z) = V \frac{Z}{H(X)} - \frac{1}{2} \frac{\partial P}{\partial X} Z (H(X) - Z), \quad (36)$$

where V is the horizontal velocity at $Z = H(X)$, given by

$$\begin{aligned} V &= \dot{X}_0 + \Omega [1 + \eta(\theta - \phi)] \cos \theta, \\ &= \dot{X}_0 + \Omega \sqrt{(1 + \eta(\theta - \phi))^2 - X^2}, \\ &= \dot{X}_0 + (Z_0 - H) \Omega \end{aligned} \quad (37)$$

By integrating the continuity equation in Z , we find that

$$\frac{\partial}{\partial X} \left(\int_0^{H(X)} U dZ \right) - \frac{\partial H}{\partial X} V + W = 0, \quad (38)$$

where W is the vertical velocity at $Z = H(X)$, namely

$$\begin{aligned} W &= \dot{Z}_0 + \Omega [1 + \eta(\theta - \phi)] \sin \theta, \\ &= \dot{Z}_0 + X \Omega \end{aligned} \quad (39)$$

By combining equations (36-39), we arrive at the Reynolds equation

$$\frac{\partial}{\partial X} \left[-\frac{H^3}{12} \frac{\partial P}{\partial X} - \frac{1}{2} H \dot{X}_0 + X \dot{Z}_0 + \frac{1}{2} (X^2 - Z_0 H) \Omega \right] = 0. \quad (40)$$

After taking a first integral of (40), we use the boundary condition $P(\pm 1) = \int_{-1}^1 \partial_X P dX = 0$, to determine the constant of integration and express the pressure gradient in the form

$$\frac{\partial P}{\partial X} = \frac{6}{H^3 I_{03}} \left[\dot{X}_0 (I_{02} - H I_{03}) + 2 \dot{Z}_0 (X I_{03} - I_{13}) + \Omega (X^2 I_{03} - I_{23} + Z_0 I_{02} - Z_0 H I_{03}) \right], \quad (41)$$

where the constants I_{mn} are defined by the integrals

$$I_{mn} = \int_{-1}^1 \frac{X^m}{H(X)^n} dX. \quad (42)$$

Having obtained expressions for the velocity field (36) and pressure gradient (41), we can proceed to find the force and torque on the rough cylinder. The force is given by

$$\begin{aligned} \frac{\mathbf{F}}{M'g} &= \frac{(F_X, F_Z)}{M'g} = \int \sigma \cdot \mathbf{n} dS \\ &= \int_{-1}^1 \left(\left\{ -P + \frac{\partial U}{\partial X} \right\} \frac{\partial H}{\partial X} - \frac{1}{2} \frac{\partial U}{\partial Z}, P \right) dX + \mathcal{O} \left(\left[\frac{\partial H}{\partial X} \right]^2 \right) \\ &= \int_{-1}^1 \left(\frac{3H}{4} \frac{\partial P}{\partial X} - \frac{Z_0}{2H} \dot{X}_0 + \frac{\Omega}{2}, -X \frac{\partial P}{\partial X} \right) dX + \mathcal{O} \left(\left[\frac{\partial H}{\partial X} \right]^2 \right). \end{aligned} \quad (43)$$

We neglect the higher order terms in accordance with the lubrication approximation. Substituting for the pressure gradient using (41), we can express the force on the cylinder as

$$\frac{\mathbf{F}}{6M'g} = (M_{11} \dot{X}_0 + M_{12} \dot{Z}_0 + M_{13} \Omega, M_{21} \dot{X}_0 + M_{22} \dot{Z}_0 + M_{23} \Omega), \quad (44)$$

where

$$M_{11} = \frac{3}{4} (I_{02}^2 - \frac{2}{3} I_{01} I_{03}) / I_{03}, \quad (45)$$

$$M_{12} = \frac{3}{2} (I_{03} I_{12} - I_{13} I_{02}) / I_{03}, \quad (46)$$

$$M_{13} = \left[\frac{1}{6} I_{03} + \frac{3}{4} (I_{03} I_{22} - I_{23} I_{02}) + \frac{3}{4} Z_0 (I_{02}^2 - \frac{8}{9} I_{01} I_{03}) \right] / I_{03}, \quad (47)$$

$$M_{21} = (I_{02} I_{13} - I_{03} I_{12}) / I_{03}, \quad (48)$$

$$M_{22} = 2(I_{03} I_{23} - I_{13}^2) / I_{03}, \quad (49)$$

$$M_{33} = [I_{03} I_{33} - I_{13} I_{23} - Z_0 (I_{03} I_{12} - I_{02} I_{13})] / I_{03}. \quad (50)$$

The torque about the centre of the unperturbed cylinder is given by

$$\begin{aligned} \frac{G}{M'gR} &= \int (1 + \eta) \mathbf{t} \cdot \sigma \cdot \mathbf{n} dS, \\ &= \int_{-1}^1 \frac{1 + \eta(\theta - \phi)}{2} \frac{\partial U}{\partial Z} dX, \\ &= \frac{1}{4} \int_{-1}^1 \sqrt{X^2 + (Z_0 - H)^2} \left\{ H \frac{\partial P}{\partial X} + 2 \frac{\dot{X}_0}{H} + \frac{2(Z_0 - H)\Omega}{H} \right\} dX. \end{aligned} \quad (51)$$

In obtaining (51), we have used the expression for H (35) to remove the explicit dependence on the roughness η . After substituting for the pressure gradient using (41) and evaluating the integrals, we find the torque can be written in the form

$$\frac{G}{M'gR} = M_{31}\dot{X}_0 + M_{32}\dot{Z}_0 + M_{33}\Omega, \quad (52)$$

where

$$M_{31} = (\frac{3}{2}I_{02}J_{02} - I_{03}J_{01})/I_{03}, \quad (53)$$

$$M_{32} = 3(I_{03}J_{12} - I_{13}J_{02})/I_{03}, \quad (54)$$

$$M_{33} = [-\frac{1}{2}I_{03}J_{00} + \frac{3}{2}(I_{03}J_{22} - I_{23}J_{02}) - Z_0(I_{03}J_{01} - \frac{3}{2}I_{02}J_{02})]/I_{03}. \quad (55)$$

Here, the coefficients J_{mn} are defined by

$$J_{mn} = \int_{-1}^1 \sqrt{X^2 + (H - Z_0)^2} \frac{X^m}{H^n} dX. \quad (56)$$

This completes our derivation of the resistance matrix \mathbf{M} , which is defined by its components (45-50) and (53-55). Given the force and torque acting on the cylinder, we can invert M to find the velocity and angular velocity, provided that \mathbf{M} is nonsingular.

2.2.2 Free cylinder

For the physically important problem of settling along an inclined plane under gravity, we assume that the roughness is sufficiently small that we may neglect any gravitational torque. We therefore set $G = 0$, $F_X = M'g \sin \alpha$, and $F_Z = M'g \cos \alpha$ to find that

$$\dot{X}_0 = \frac{M'g}{\det(M)} [\sin \alpha (M_{22}M_{33} - M_{12}M_{21}) + \cos \alpha (M_{13}M_{32} - M_{12}M_{23})], \quad (57)$$

$$\dot{Z}_0 = \frac{M'g}{\det(M)} [\sin \alpha (M_{23}M_{31} - M_{21}M_{33}) + \cos \alpha (M_{11}M_{13} - M_{31}M_{13})], \quad (58)$$

$$\Omega = \frac{M'g}{\det(M)} [\sin \alpha (M_{21}M_{32} - M_{31}M_{22}) + \cos \alpha (M_{13}M_{21} - M_{11}M_{23})]. \quad (59)$$

It should be noted that all of the coefficients M_{ij} depend upon two parameters, namely the current height Z_0 and the current orientation ϕ . The equations (57-59) therefore define an autonomous ordinary differential equation for the position and orientation of the cylinder as a function of time. Note also that the evolution of the downslope position is slaved to the height and orientation.

In order to investigate the behaviour of the system (57-59), we consider a sample roughness function in the form of a sawtooth with m teeth – i.e. a periodic function

$$\eta(\theta) = \eta \left(\theta + \frac{2n\pi}{m} \right) = A\theta, \quad -\frac{\pi}{m} < \theta < \frac{\pi}{m}, \quad n = \pm 0, \pm 1, \dots, \pm(m-1). \quad (60)$$

The results presented here are for $m = 10$, but we shall discuss other values of m at the end of this subsection.

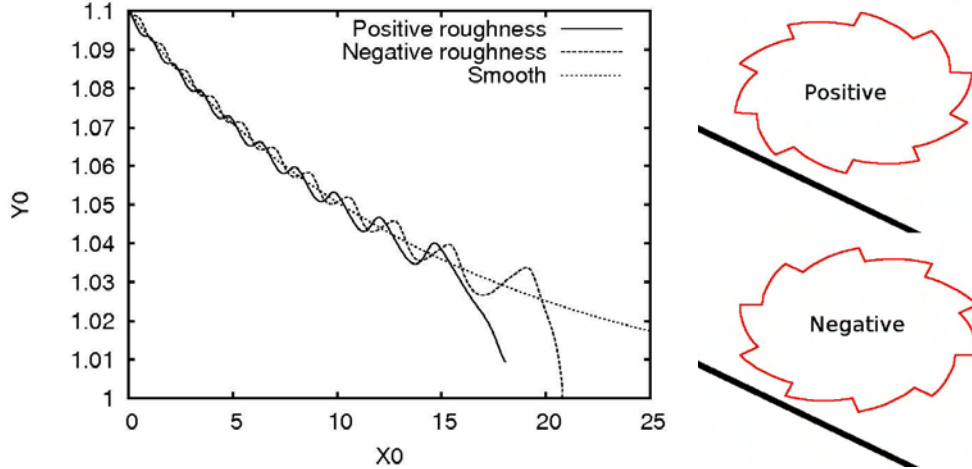


Figure 8: Trajectory of the centre of mass (X_0, Z_0) of a rough cylinder falling freely down a plane inclined at angle $\alpha = \pi/6$. Positive and negative roughness corresponds to the sawtooth roughness function (60) with $m = 10$ and $A = \pm 0.01$ respectively.

Figure 8 shows the trajectories described by the centre of the cylinder (X_0, Z_0) in typical numerical simulations. For comparison, we also show the trajectory of a smooth cylinder under the same conditions. We see that the paths of the rough and smooth cylinders are largely similar. When far from the inclined plane, the path of the rough cylinder oscillates about the path of the smooth cylinder. For ‘positive’ roughness, periods of gradual ascent are interspersed by shorter periods of relatively rapid descent. For ‘negative’ roughness, we observe short periods of ascent interrupting longer periods of gradual descent. In both cases, however, there is a net decrease in height over these two intervals.

Given the absence of such oscillations in the case of a smooth cylinder, the only cause can be the addition of the symmetry-breaking roughness to the cylinder. In fact, the longer periods of gradual ascent or enhanced descent can be explained by the sawteeth acting as Reynolds bearings as illustrated in Figure 9. In each case, gravity forces the cylinder to move downslope. For negative roughness, the local geometry of a tooth moving past the inclined plane resembles that of a Reynolds bearing (Figure 9a), where hydrodynamical forces in the converging channel generate a net lift force on the cylinder. For positive roughness, the situation is reversed (Figure 9b), and the net hydrodynamical force on the cylinder is directed towards the plane. When added to the resistive pressure force that comes from squeezing out fluid from beneath the cylinder, this bearing effect can either oppose or enhance the settling of the cylinder due to gravity.

However, gravity also exerts a torque on the cylinder, forcing it to rotate so that the Reynolds bearing/antibearing geometry changes to one in which a sharp tooth is near the base of the cylinder. This gives rise to the shorter periods of motion in the opposite direction. In the case of positive roughness, the passing of the tooth abruptly increases the thickness of the gap between the cylinder and the plane, reducing the resistive forces and allowing for an increased settling velocity. For negative roughness, the sharp decrease in the

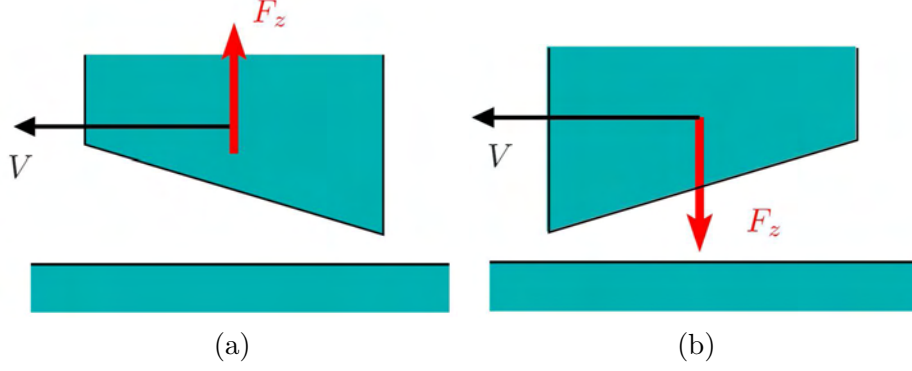


Figure 9: Cartoon of Reynolds bearings. (a) If the fluid moves towards a narrowing gap, a lift force is generated. (b) If fluid moves towards an expanding gap, a suction force arises.

gap thickness results in a short-lived upward force, and allows the cylinder to rise slightly before settling once again into the falling part of the cycle.

Whether the roughness is in the form of a positive or negative sawtooth, the end result is essentially the same. When the cylinder comes close enough to the plane, the gravitational torque is insufficient to keep the cylinder rolling downslope, and it gets stuck with a sharp point near its base. The numerical results seem to suggest that contact in finite time then follows (though it is difficult to distinguish finite-time contact from a finite-time approach predicted by a numerical scheme with too large a timestep). However, we could believe, given the results of §2.1, that the sharp corner does allow finite-time contact.

The above discussion applied for $m = 10$. If we change m , the number of teeth on the rough cylinder, the results remain qualitatively unchanged. In each case, it seems that the the cylinder will reach contact in finite time, and its trajectory oscillates around that of a smooth cylinder until it comes sufficiently close to the inclined plane, at which point rotation is arrested and the sharp corner can begin to effect finite-time settling.

After the preparation of an early form of this report, a paper by Zhao *et al.*[6] was brought to our attention. In this paper, the authors investigate experimentally the rolling of a smooth sphere down an inclined plane of carefully-controlled roughness. They look in detail at the dynamics of the sphere as it passes over an asperity, and use a similar lubrication model to describe them. We direct the interested reader to [6] for a much more detailed and polished version of the above analysis.

2.2.3 Driven cylinder

As a brief aside, we draw attention to the dynamics of a roughened cylinder forced to rotate at constant (dimensionless) angular velocity Ω_0 , but that is otherwise free to move. In order to allow for comparison with the case of a free cylinder, we use the same definitions of positive and negative roughness (i.e. the sawtooth roughness of (60) with $m = 10$ and $A = \pm 0.1$). We present here results for zero inclination angle $\alpha = 0$, but will mention the effect of varying α at the end of this subsection. The restriction of the free cylinder problem to a driven problem can be achieved under the previous resistance matrix framework by

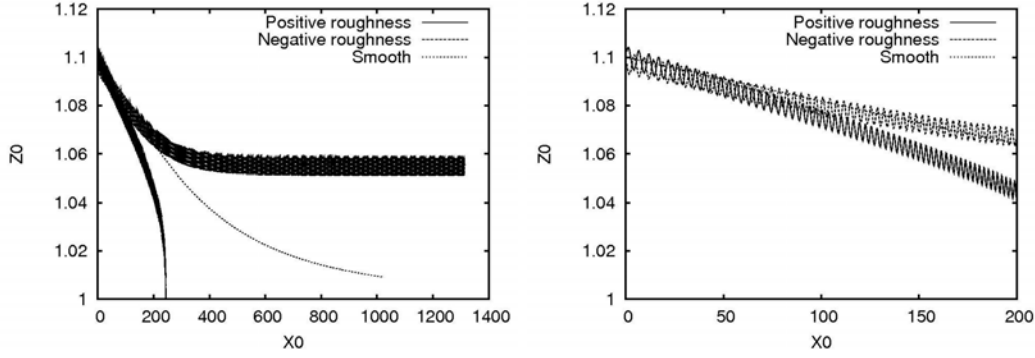


Figure 10: Trajectories of a driven cylinder falling down an inclined plane. Positive and negative roughness correspond to the sawtooth roughness (60) with $m = 10$ and amplitude $A = \pm 0.1$. The plane beneath the cylinder is horizontal – $\alpha = 0$. The trajectory of a smooth cylinder is shown for comparison. The second graph shows a magnification of the initial region $0 < X_0 < 200$, showing detail of the ratchet-like motion.

setting $M_{31} = M_{32} = 0$, $M_{33} = 1$, and $G = \Omega_0$. Figure 10 shows the trajectories followed by the centre of mass of the cylinder in the cases of positive and negative roughness, where the cylinder is driven in a clockwise sense.

In the case of positive roughness, we see that the cylinder rapidly descends towards the horizontal plane, and appears (numerically) to make contact in finite time. This can be explained in a similar manner to the apparent finite-time contact observed for the free cylinder. In this case, however, the settling is greatly accelerated by the rotation of the cylinder, which generates a strong suction force via Reynolds anti-bearing action (Figure 9b).

In the case of negative roughness, however, we observe that, following an initial period where the cylinder settles at a (time-averaged) rate similar to that of a smooth cylinder, the vertical position of the centre of mass, Z_0 tends to a stable periodic oscillation about the value $Z_0 \approx 1.05$. In other words, the cylinder will remain suspended above the plane indefinitely, as long as it continues to be driven to rotate. The existence of this suspended state is quite robust to variation of the rotation rate and roughness parameters, though the limiting mean value of Z_0 decreases as the rotation rate, roughness amplitude, and number of teeth decrease.

In both of the above cases, looking more closely at the trajectories reveals a ratchet-like behaviour, similar to the case of a free rough cylinder, where longer periods of gradual rise or fall are interspersed with short periods of motion in the opposite direction. However, in this case, it is the act of driving the cylinder that produces horizontal motion, rather than gravity. If the cylinder rotates fast enough in the appropriate direction for the roughness, then it can provide sufficient lift (during those intervals when we have a Reynold bearing geometry) for the net change in height during the passing of one tooth to be zero at some critical height. This gives rise to the periodic oscillation observed here for negative roughness.

Varying the inclination angle does little to change the qualitative dynamics of the driven system. A fairly coarse search of the parameter space indicates that the suspended state

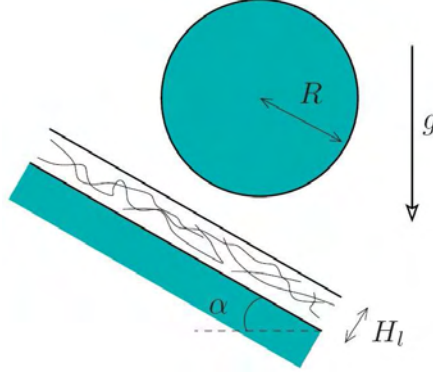


Figure 11: Diagram and notation for elastic lubrication problem

still exists for any angle, provided that the sense of the rotation is correct. For a large enough angle, it may be the case that the downslope component of force can balance the upslope component of the viscous drag due to rotation, and one could find a system in which both X_0 and Z_0 exhibit periodic oscillations. The search for such a state did not, however, form a part of the work carried out during the GFD Program.

3 Elasticity

The infinite contact time predicted by a simple lubrication model requires the appearance of very large fluid pressures when the gap between the falling object and the substrate are very small. It is quite possible that these large pressures could cause some small deformation of the solid surfaces, which were previously assumed to be rigid. It is not immediately obvious whether elastic effects should aid or oppose settling. In this section, we develop ideas proposed by Skotheim & Mahadevan [4] about lubrication films between elastic layers, and attempt to apply them to settling problems.

3.1 Lubrication with elastic layers

We consider a situation such as that shown in Figure 11. A smooth cylinder of radius R is settling under gravity towards an smooth plane, inclined at angle α to the horizontal, that is coated with a thin layer of thickness H_l of an elastic material. We define a local Cartesian coordinate system (x, z) to be aligned with the inclined plane, and take $x = 0$ to pass through the centre of the cylinder.

We assume that the elastic material has a linear, isotropic stress-strain relationship, with stress tensor

$$\sigma_e = G(\nabla \mathbf{u} + (\nabla \mathbf{u})^T) + \lambda(\nabla \cdot \mathbf{u})\mathbf{I}, \quad (61)$$

where \mathbf{u} is the local displacement vector, \mathbf{I} is the identity tensor, and G and λ are Lamé coefficients describing the elasticity of the medium. The elastic problem to be solved,

assuming a quasi-steady state, is

$$\nabla \cdot \sigma_e = 0 \quad (62)$$

$$\mathbf{u}(X, -H_l) = \mathbf{0} \quad (63)$$

$$\sigma_e \cdot \mathbf{n} = -p\mathbf{n} \quad \text{at } z = u_z(x, 0) \quad (64)$$

which corresponds to a local force balance, with no displacement allowed on the lower boundary, and continuity of stress on the surface in contact with the fluid. If we assume that the layer is thin compared to the horizontal extent of the lubrication film, then we can proceed in the same manner as fluid lubrication, by neglecting horizontal variations when determining the displacement field. We therefore approximate (62) by

$$(2G + \lambda) \frac{\partial^2 u_z}{\partial z^2} = 0,$$

giving us the approximate solution

$$u_z(x, z) \approx -\frac{p(x)}{2G + \lambda}(z + H_l). \quad (65)$$

The size of the fluid-filled gap between the settling object and the elastic layer where the local pressure is $p(x)$ is given by

$$h(x) = h_0(x) + \frac{H_l}{2G + \lambda} p(x),$$

where $h_0(x)$ is the undeformed gap thickness. For the case of a settling cylinder, we approximate this by a parabola, leading to

$$h(x) = \epsilon R + \frac{x^2}{2R} + \frac{H_l}{2G + \lambda} p(x), \quad (66)$$

where ϵR is the minimum separation between the cylinder and the line $z = 0$.

The equations governing the fluid flow can be determined exactly as for a rigid cylinder approaching a rigid substrate. For a cylinder translating with velocity (\dot{x}_0, \dot{z}_0) , we find the Reynolds equation

$$\frac{\partial}{\partial x} \left[-\frac{h^3}{12\mu} \frac{\partial p}{\partial x} - \frac{h}{2} \dot{x}_0 + x \dot{z}_0 \right] = 0. \quad (67)$$

The novel feature in this elastic problem is the feedback between pressure and gap thickness, as established by (66). Some consequences of this are discussed in the next few sections.

3.2 Pure shearing

The first use of the preceding theory was made by Skotheim & Mahadevan [4], who considered the case of a cylinder travelling at a fixed height past an elastic-coated plane with speed V , as illustrated by Figure 12a. In the absence of an elastic layer, the symmetry of the translating cylinder gives rise to a perfectly symmetric pressure gradient (Figure 12b), which cannot generate lift. However, the elastic layer can be deformed by the relative pressures, and is squeezed down ahead of and beneath the cylinder, as shown in Figure 12c.

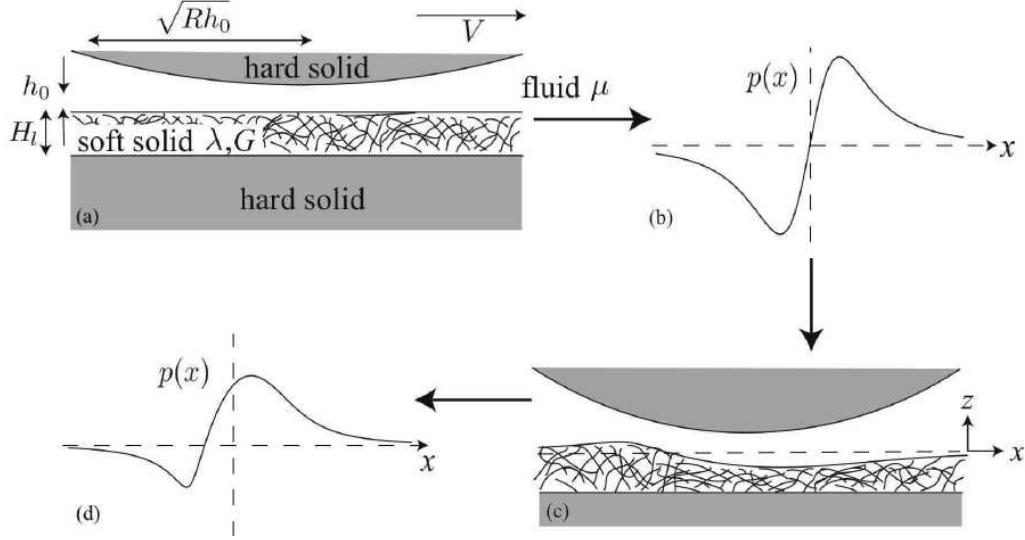


Figure 12: Illustration of how the presence of an elastic layer can generate lift. (a) A rigid object translates parallel to an elastic layer. (b) The symmetric pressure distribution that arises without elasticity. (c) Deformation of the elastic layer due to high and low relative pressures. (d) The resulting asymmetric pressure distribution. Figure copied from Skotheim & Mahadevan (2004) [4].

This in turn makes the pressure gradient asymmetric (Figure 12d), and can generate a net upward force. The mathematical details of this problem are given in full in [4], wherein the authors also consider how the lift force scales with the separation of the object from the elastic material for a range of different physical situations, including either solid object being entirely elastic and the cylinder being an elastic shell. In the context of this report, we note merely the existence of a mechanism by which the presence of elasticity in the solid bodies can generate lift, and therefore oppose the settling of an object, provided that there is some force driving the object horizontally.

3.3 Pure settling

In the absence of any horizontal forcing ($\alpha = 0$ in Figure 11), we find a situation like that shown in Figure 13. In this case, we expect that the pressure is an even function of x , and the Reynolds equation (67) can be integrated once to find that

$$\frac{\partial p}{\partial x} = \frac{12\mu R x \dot{\epsilon}}{h^3(x)}, \quad (68)$$

the constant of integration having been eliminated by integrating over $-\infty < x < \infty$ and applying the boundary conditions $p(\pm\infty) = 0$. At this point, we nondimensionalise the problem by scaling all lengths with R , time with $\mu R/M'g$, and pressure with $M'g/R$, where $M'g$ is the effective weight per unit length of the cylinder. Substituting for the approximate gap width (66), and balancing the vertical resistive force with the effective weight of the

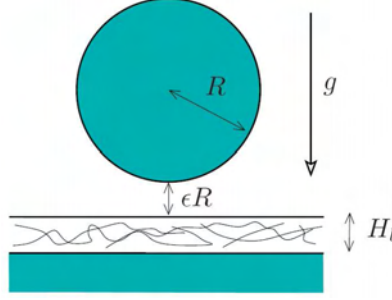


Figure 13: Diagram and notation for a vertically-settling cylinder onto an elastic layer.

cylinder, we find the the dimensionless pressure P must be a solution to the following problem

$$\frac{dP}{dX} = \frac{12\epsilon X}{(\epsilon + \frac{1}{2}X^2 + \gamma P)^3}, \quad (69)$$

$$P(\pm\infty) = 0, \quad (70)$$

$$\int_{-\infty}^{\infty} P dX = 1, \quad (71)$$

where

$$\gamma = \frac{M'gH_l}{(2G + \lambda)R^2}. \quad (72)$$

In order to solve this problem, we define the new variables ψ and ξ by

$$\xi = \epsilon^{-1/2}X, \quad \text{and} \quad \frac{d\psi}{dX} = \frac{\gamma P}{\epsilon}.$$

With an appropriate choice of the arbitrary constant allowed by our choice of ϕ , we transform the problem (69-71) into the second-order boundary value problem

$$\psi'' = \frac{N\xi}{(1 + \frac{1}{2}\xi^2 + \psi')^3}, \quad (73)$$

$$\psi'(\pm\infty) = 0, \quad (74)$$

$$\psi(\pm\infty) = \pm \frac{\gamma}{2\epsilon^{3/2}}. \quad (75)$$

Here $N(= 12\gamma\epsilon/\epsilon^3)$ may be thought of as an eigenvalue for the ordinary differential equation (73) with boundary conditions (74,75). This problem can be easily solved numerically for all values of the parameter δ , resulting in the relationship between the settling speed $\dot{\epsilon}$ and the separation ϵ shown in Figure 14. Importantly, this seems to indicate that the settling velocity tends to a nonzero limit as $\epsilon \rightarrow 0$, suggesting that the cylinder *will* settle in finite time. A physical explanation for this is as follows. Deformation of the elastic film increases the mean fluid-filled gap thickness, which in turn has the result of reducing the pressure gradient (recall that the pressure gradient is proportional to h^{-3}). This results in a net

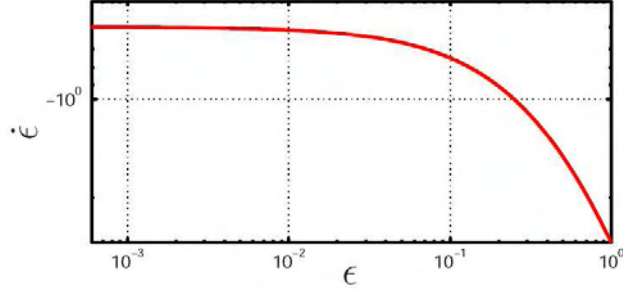


Figure 14: Graph showing the settling velocity $\dot{\epsilon}$ as a function of separation ϵ for vertical settling on to an elastic layer. Note that $\dot{\epsilon}$ tends to a finite nonzero limit as $\epsilon \rightarrow 0$.

decrease in the upward resistive force, so that $F_Z = \text{ord}(1)$ as $\epsilon \rightarrow 0$. We can understand this asymptotic behaviour analytically by defining scaled variables (ϕ, ζ) such that

$$\psi(\xi) = \frac{\gamma}{2\epsilon^{3/2}}\phi(\zeta) \quad \text{and} \quad \xi = \left(\frac{2\epsilon^{3/2}}{\gamma}\right)^{1/3} \zeta.$$

The equations (73-75) may then be expressed as

$$\frac{d^2\phi}{d\zeta^2} = N \left(\frac{2\epsilon^{3/2}}{\gamma}\right)^2 \frac{\zeta}{(\frac{1}{2}\zeta^2 + \frac{d\phi}{d\zeta} + \mathcal{O}(\epsilon))^3}, \quad (76)$$

$$\frac{d\phi}{d\zeta}(\pm\infty) = 0, \quad (77)$$

$$\phi(\pm\infty) = \pm 1.. \quad (78)$$

In the limit $\epsilon \rightarrow 0$, this boundary value problem can only admit solutions if the right-hand side of (73) is of order unity. That is

$$N\epsilon^3 = \text{ord}(1) \Rightarrow \dot{\epsilon} = \text{ord}(1),$$

as observed in the results of the numerical calculation described above.

This scaling for the settling velocity suggests that the addition of elasticity should allow contact in finite time, because $\epsilon \propto (t - t_0)$ for small ϵ . However, one must remember that ϵ is the height of the cylinder above the *undeformed* elastic layer. After passing through $\epsilon = 0$, the layer will simply continue to deform at a rate that maintains the $\text{ord}(1)$ scaling for $\dot{\epsilon}$. There will come a point, of course, when the elastic layer can deform no longer, either due to nonlinear elastic effects, or the fact that deformation predicted by this simple model will exceed the thickness of the layer. At this point, we should probably expect to see the familiar $\epsilon \propto t^{-2}$ scaling for the approach of two rigid objects reappear. Our conclusion is therefore that elasticity can accelerate contact at first, but should not be able to create bona fide contact between the objects in finite time.

3.4 Arbitrary inclined plane

We now return to the situation depicted in Figure 11, where a rigid cylinder settles down an inclined plane coated with an elastic layer. In this case, the downslope component of gravity acts to drag the cylinder downhill, giving rise to a lift force as described in §3.2. Meanwhile, the perpendicular component acts to force the cylinder to settle. It is not immediately obvious which of these two effects will dominate the dynamics of the cylinder, so we shall investigate the system numerically in this section.

For simplicity, we assume that the cylinder does not rotate, and just translates with velocity (\dot{x}_0, \dot{z}_0) . Using the same scalings for pressure, length, time and velocity as in §3.3, the Reynolds equation (67) for the pressure gradient beneath the cylinder becomes

$$\frac{\partial}{\partial X} \left[-\frac{H^3}{12} \frac{\partial P}{\partial X} - \frac{H}{2} \dot{X}_0 + X \dot{Z}_0 \right] = 0 \quad (79)$$

Vertical force balance, taken up to second order in $\partial H / \partial X$, once again requires that we equate the pressure force with the plane-perpendicular component of the weight,

$$\int_{-\infty}^{\infty} P \, dX = \cos \alpha \quad (80)$$

whilst the downslope force balance becomes, after a little algebra

$$\int_{-\infty}^{\infty} \left\{ \frac{\dot{X}_0}{2H} + \frac{3H}{4} \frac{\partial P}{\partial X} \right\} dX = -\sin \alpha. \quad (81)$$

Along with the boundary conditions $P(\pm\infty) = 0$, equations (79-81) define a boundary value problem with two eigenvalues, namely \dot{X}_0 and \dot{Z}_0 . This problem can be solved numerically by defining new dependent variables Ψ , Φ , and independent variable ξ by

$$\frac{\partial P}{\partial X} = \frac{\epsilon^{1/2}}{\gamma} \Psi'(\xi), \quad \Phi'(\xi) = \tilde{H} \Psi'' + \frac{V}{9\tilde{H}}, \quad x = \epsilon^{1/2} \xi,$$

so chosen to turn the integral constraints (80) and (81) into boundary conditions on Ψ and Φ , respectively. The scaled height \tilde{H} is given by

$$\tilde{H} = 1 + \frac{1}{2} \xi^2 + \gamma P.$$

and

$$V = \frac{6\gamma \dot{X}_0}{\epsilon^{5/2}}.$$

With this transformation, we must solve the third-order boundary value problem

$$\Psi'' = \frac{A + N\xi}{\tilde{H}^3} - \frac{V}{\tilde{H}^2}, \quad (82)$$

$$\Phi' = \tilde{H} \Psi'' + \frac{V}{9\tilde{H}}, \quad (83)$$

$$\Psi(\pm\infty) = \Psi'(-\infty) = \Phi(-\infty) = 0, \quad (84)$$

$$\Psi(+\infty) = \frac{\gamma \cos \alpha}{\epsilon^{3/2}}, \quad (85)$$

$$\Phi(+\infty) = -\frac{4\gamma \sin \alpha}{3} \left(\frac{\gamma \cos \alpha}{\epsilon^{3/2}} \right)^{4/3}. \quad (86)$$

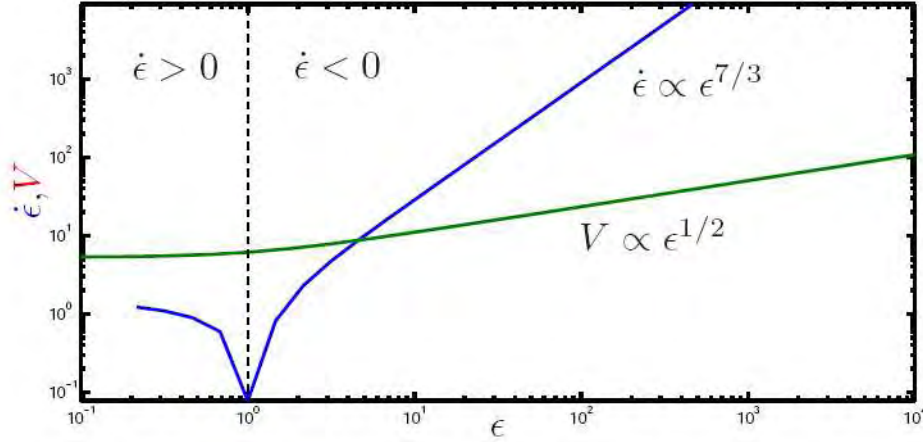


Figure 15: Log-log plot of numerical results for the eigenvalues N (in blue, compensated by ϵ^3) and V (in red, compensated by $\epsilon^{3/2}$), as a function of the separation ϵ for a cylinder falling down an elastic-coated inclined plane. Note the change in sign of N (hence \dot{Z}_0) in the vicinity of $\epsilon = 1$.

In equation (82), A is the constant of integration introduced when taking a first integral of the Reynolds equation (79). This must also be determined as an eigenvalue of the third-order problem, along with the quantities N and V , which relate to the settling and downslope velocities, respectively. In this problem, $N = 12\gamma\dot{Z}_0/\epsilon^3$, as in §3.3.

Solving equations (82-86) numerically, one typically finds results like those shown in Figure 15. The very important feature to note is that the eigenvalue N changes sign somewhere near $\epsilon = 1$. This suggests the existence of a (stable) steady state, where the cylinder falls down the plane at a constant velocity and constant separation from the plane. In this situation, the settling force exactly balances the lift force due to the downslope motion. Note that, unlike the driven cylinder of §2.2.3, this 'gliding' stable state exists without the need for any external forcing beyond gravity. We have also noted the apparent power-law scalings of N and V for large separations ϵ , but have not yet made an attempt to explain them physically or mathematically.

4 Compressibility

A final factor that may play a role in settling dynamics is the compressibility of the fluid film. One might expect that the high lubrication pressures could be somewhat alleviated by allowing the density of the fluid to change. With this in mind, we consider once again the problem of a settling cylinder, as illustrated in Figure 1, but this time allow for compressible effects in the fluid film.

The lubrication approximation to the equations of conservation of mass and momentum

(based on the Stokes equations) are

$$-\frac{\partial p}{\partial x} + \mu \frac{\partial^2 u}{\partial x^2} = 0, \quad (87)$$

$$\frac{\partial p}{\partial z} = 0, \quad (88)$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) + \frac{\partial}{\partial z}(\rho z) = 0. \quad (89)$$

These equations must be supplemented by an equation of state. In the case of an adiabatic gas, pressure and density are related by

$$p = K \rho^\gamma. \quad (90)$$

We also suppose that the kinematic viscosity varies with the local pressure (hence density) according to

$$\nu = L p^m. \quad (91)$$

As in the case of an incompressible fluid, we can integrate the momentum equations (87,88) to obtain a parabolic velocity profile

$$u = -\frac{\partial_x p}{2\mu} z(h - z) \quad (92)$$

and integrate the mass conservation equation (89) in the vertical direction to find the compressible Reynolds equation

$$\frac{\partial}{\partial t}(\rho h) + \frac{\partial}{\partial x} \left(-\frac{h^3}{12\nu} \frac{\partial p}{\partial x} \right) = 0. \quad (93)$$

This equation, together with the relations (90,91) can be solved for the pressure p , subject to boundary conditions $p(\pm\infty) = 0$. Finally, we use instantaneous force balance

$$\int p \, dx = M g' \quad (94)$$

to determine the settling velocity as a function of the current shape and size of the fluid-filled gap, as in the previous sections.

Though the Reynolds equation (93) may be integrated numerically, we shall only note the existence of a late-time similarity solution. We nondimensionalise the problem by writing

$$x = RX, \quad h = RH, \quad p = \frac{M g'}{R} P, \quad t = 12L \left(\frac{M g'}{R} \right)^{1-m-1/\gamma}.$$

After substituting for the density and kinematic viscosity, we arrive at the partial differential equation

$$\frac{\partial}{\partial T} \left(H P^{1/\gamma} \right) - \frac{\partial}{\partial X} \left(\frac{H^3}{P^m} \frac{\partial P}{\partial X} \right) = 0, \quad (95)$$

$$P(\pm\infty) = 0, \quad (96)$$

$$\int P \, dX = 1 \quad (97)$$

Noting that we may approximate the gap thickness by a parabola

$$H(X, T) \approx \epsilon(T) + \frac{1}{2}X^2 = \epsilon(T) \left[1 + \frac{1}{2} \left(\frac{X}{\sqrt{\epsilon}} \right)^2 \right] =: \epsilon \tilde{H}(\xi), \quad (98)$$

we aim to find a large- T similarity solution in terms of the similarity variable $\xi = X/\sqrt{\epsilon}$. By seeking a solution of the form

$$P(X, T) = T^\lambda f(\xi),$$

we can transform the problem (95-97) into an ordinary differential equation for in ξ only provided that

$$\frac{\dot{\epsilon}}{\epsilon^2} \propto T^{\lambda(1-m-1\gamma)}$$

and

$$t^\lambda \sqrt{\epsilon} \propto 1.$$

By combining these two relationships, we find that the long-time scaling of the minimum gap width is given by

$$\epsilon \propto T^{-2/(1+m+1\gamma)}. \quad (99)$$

In the case of a fluid with constant *dynamic* viscosity, we require that $m = -1/\gamma$, so we recover the $\epsilon \propto T^{-2}$ scaling obtained in the incompressible case, regardless of the value of γ . If instead we assume that the *kinematic* viscosity remains constant, then $m = 0$, and the exponent in (99) is strictly greater than -2 for all (positive) values of γ . In this latter case, the settling velocity is decreased due to the increase in dynamic viscosity that arises when the fluid is compressed. The similarity scaling (99) has been verified against a direct integration of the partial differential equation (95) in the case of both constant dynamic viscosity and constant kinematic viscosity.

In summary, compressibility does not appear to provide a means by which a contact can be made in finite time. Indeed, it is difficult to improve on the $\epsilon \sim T^{-2}$ approach obtained in the incompressible problem described in §1. Whilst compressing the fluid can alleviate some of the pressure beneath the settling object, there is an associated increase in dynamic viscosity that provides a greater resistance to flow. This, in turn, increases the pressure beneath the object, slowing its approach.

5 Conclusions and future work

At the outset, the intention of this project was to consider whether simple hydrodynamical effects could resolve the problem of infinite contact time for two solid objects coming together in a viscous fluid. We chose to avoid a discussion of short-ranged effects such as Van der Waals forces, in favour of the effects of the shape and elasticity of the objects, and compressibility of the fluid. Of these three effects, it seems that only shape, in the form of sharp apertures, can resolve this long-standing problem.

We have shown that an object with discontinuous first derivative will settle towards a flat surface in finite time by considering the model problem of a sedimenting triangular

wedge. However, we considered a constrained problem, in which the wedge was not allowed to change its orientation relative to the ground. Given that the wedge would be unstable to toppling in one direction or the other, it would be interesting to examine the motion of the wedge in the unconstrained problem, where the wedge is entirely free to move. Would contact still occur in finite time, or would the wedge orient itself so that one of its edges becomes parallel to the ground, thus leading to an infinite contact time? Many of the methods discussed in §2.1 could be used to model the falling wedge, including both the outer Stokes approximation and the numerical scheme presented therein. This remains as a strong candidate for future work.

During the GFD Program, we expended some time and effort on modelling the rolling of a rough cylinder next to a horizontal surface. We found that the presence of roughness can, in certain cases, either result in contact in finite time or no contact at all. While a more thorough investigation of how asymmetric roughness must be in order to cause these effects would be an interesting avenue for future study, our key results were essentially pre-empted by Zhao *et al.* [6]. Once again, we direct the reader to this paper for a more developed version of the results of §2.2.

The effects of material elasticity and fluid compressibility were investigated briefly, but it was found that neither could provide a mechanism by which contact could be achieved in finite time. However, our analysis of the problem with elastic boundaries suggested that it is possible to avoid contact entirely, provided that there is a sufficiently large force acting perpendicular to the line of closest approach of the two objects. The physics behind this effect is well described by the work of Skotheim & Mahadevan [4].

In summary, it appears that only surface roughness can predict finite time contact without resorting to non-hydrodynamical effects. Even then asperities must be perfectly sharp in order to make contact. For real surfaces, of course, no such sharp corners exists on (and above) the molecular scale, so it is the opinion of the author that one will be drawn inevitably towards the consideration of long range forces, and the dewetting of the by then very thin film between the two objects. This problem remains as a challenge to future researchers.

I would like to take the opportunity to thank everyone involved with this year's GFD Program for a thoroughly enjoyable summer. In particular, I should like to thank Neil for his mathematical guidance and outstanding singing ability²; George, Charlie and Phil for their patience and encouragement on the softball field; and Maha, Joe, and Bill for useful discussions. Finally, I would like to thank my fellow Fellows for being a lovely group of people to get to know over the summer, as well as being useful sounding boards for my 'obvious' mathematical deductions.

References

- [1] G. K. BATCHELOR, *An introduction to fluid mechanics*, Cambridge University Press, 1967.
- [2] D. GÉRARD-VARET AND M. HILLAIRET, *Regularity issues in the problem of fluid structure interaction*, ArXiv e-prints, 805 (2008).

²Video evidence of which may be available electronically from the author, unless Neil gets to him first.

- [3] F. PLOURABOUÉ AND M. BOEHM, *Multi-scale roughness transfer in cold metal rolling*, Tribol. Int., 32 (1999), pp. 45–57.
- [4] J. M. SKOTHEIM AND L. MAHADEVAN, *Soft lubrication: The elastohydrodynamics of nonconforming and conforming contacts*, Phys. Fluids, 17 (2005), p. 092101.
- [5] H. A. STONE, *On lubrication flows in geometries with zero local curvature*, Chem. Engng. Sci., 80 (2005), pp. 4838–4845.
- [6] Y. ZHAO, K. P. GALVIN, AND R. H. DAVIS, *Motion of a sphere down a rough plane in a viscous fluid*, Int. J. Multiphase Flow, 28 (2002), pp. 1787–1800.

Energy and dissipation in MHD systems

C. Gissinger

January 12, 2009

1 Introduction

In this work, Hartmann flow is studied in the framework of incompressible magnetohydrodynamics. In the first sections, we present simple examples of magnetic problem showing the role of the vacuum or an insulator on the energy and on the dissipation rate of a conductor. Then, we derive exact laminar solution and explicit expression for the energy stability of the Hartmann flow in contact with an insulator. We show that taking into account realistic boundary conditions for the magnetic field yields technical difficulties in the derivation of exact bounds on the dissipation rate using the background method.

2 Dissipation in a cylindrical conductor-1

In this section we study a simple situation of free decay of a magnetic field. We consider a wire of metal with conductivity σ and permeability μ . The wire is a cylinder infinite in the axial direction z and have a finite radius a . The cylinder is surrounded by vacuum (see figure ??). We suppose that there is a current \mathbf{J} only at $t = 0$ and we study the evolution of the magnetic field created by the initial current.

In this first simple problem, the initial current is a toroidal current depending only on the radial direction:

$$\mathbf{J}(\mathbf{r}) = J_\theta(r)\mathbf{e}_\theta \quad (1)$$

By symmetry and use of Ampere's law, we know that the corresponding initial magnetic field created is $B_z(t, r)\mathbf{e}_z$ inside the wire and is zero outside (for $r > a$). Thus, the governing equation are:

$$\frac{\partial \mathbf{B}}{\partial t} = \eta \Delta \mathbf{B} \quad (2)$$

which in cylindrical coordinates for the magnetic field inside the conducting domain is :

$$\frac{\partial \mathbf{B}_z}{\partial t} = \eta \left[-\frac{\partial^2 B_z}{\partial r^2} - \frac{1}{r} \frac{\partial B_z}{\partial r} \right] \quad (3)$$

where we have introduced the magnetic diffusivity $\eta = 1/(\sigma\mu)$. The magnetic field is simply $\nabla \times \mathbf{B} = 0$ for $r > a$. In this diffusive situation, the magnetic field can only decay and we suppose that $B_z(t, r) = b(r)e^{-\alpha t}$. The equation (3) then reduce to :

$$\frac{\partial^2 b}{\partial r^2} + \frac{1}{r} \frac{\partial b}{\partial r} + \frac{\alpha}{\eta} b = 0 \quad (4)$$

Using the change of variables $\rho = r\sqrt{\alpha/\eta}$, we obtain:

$$\frac{\partial^2 b(\rho)}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial b(\rho)}{\partial \rho} + b(\rho) = 0 \quad (5)$$

We recognize the bessel equation for $m = 0$. A solution of the equation is thus:

$$b = J_0\left(\sqrt{\frac{\alpha}{\eta}}r\right) \quad (6)$$

The boundary conditions for the magnetic field leads to an expression for the zeros of the bessel function:

$$B_z(a) = 0 \implies J_0\left(\sqrt{\frac{\alpha}{\eta}}a\right) = 0 \quad (7)$$

The n^{th} zero of the bessel function $J_0(\rho)$ is given by $\lambda_n = \sqrt{\frac{\alpha n}{\eta}}a$. We see that the boundary conditions yields here a discretization of the possible decay rates for the magnetic field. This is not surprising since these boundary conditions result in a confinement of the field in a finite domain. We can then expand the fields in Fourier-Bessel series for $r < a$:

$$b(r, t) = \sum_{n=0}^{\infty} A_n J_0(\lambda_n r/a) e^{-t\eta\lambda_n^2/a^2} \quad (8)$$

And simply $b(r, t) = 0$ for $r > a$. Because Bessel's equation is Hermitian, the solutions must satisfy the following orthogonality relationship :

$$\int_0^1 J_\nu(x\lambda_m^\nu) J_\nu(x\lambda_n^\nu) x dx = \frac{\delta_{mn}}{2} [J_{\nu+1}(\lambda_n^\nu)]^2 \quad (9)$$

In order to express the coefficient of the magnetic field we evaluate the integral :

$$I = \int_0^1 r dr J_0(\lambda_n r) B^o(r) \quad (10)$$

Where $B^o(r)$ represent the magnetic field at $t = 0$. Expansion of $B^o(r)$ gives:

$$I = \int_0^1 r dr J_0(\lambda_n r) \sum_{m=1}^{\infty} A_m J_0(\lambda_m r) \quad (11)$$

$$I = \sum_{m=1}^{\infty} A_m \int_0^1 r dr J_0(\lambda_n r) J_0(\lambda_m r) \quad (12)$$

$$I = \sum_{m=1}^{\infty} \frac{A_m}{2} \delta_{mn} [J_1(\lambda_m)]^2 \quad (13)$$

$$I = \frac{A_n}{2} [J_1(\lambda_n)]^2 \quad (14)$$

We then get an expression for the coefficient of the magnetic field:

$$A_n = \frac{2}{[J_1(\lambda_n)]^2} \int_0^1 r dr J_0(\lambda_n r) B^o(r) \quad (15)$$

The decomposition of the magnetic field on the basis of Bessel functions with well defined coefficients provide a natural way for evaluating the magnetic energy in the volume. Indeed the Parseval theorem applied to Bessel-Fourier series gives:

$$\epsilon = \int dV [\mathbf{B}(\mathbf{r}, \mathbf{t})]^2 = \sum_{n=0}^{\infty} |A_n e^{-t\eta\lambda_n^2/a^2}|^2 = \sum_{n=0}^{\infty} |A_n|^2 e^{-2t\eta\lambda_n^2/a^2} \quad (16)$$

This yields to a negative time variation of the energy:

$$\dot{\epsilon} = -2\eta \sum_{n=0}^{\infty} \lambda_n^2 |A_n|^2 e^{-2t\eta\lambda_n^2/a^2} \quad (17)$$

Finally, we use the fact that all the zeros of J_0 are greater or equal to the first zero λ_1 of the Bessel function:

$$\dot{\epsilon} = -2\eta\lambda_1^2/a^2 \sum_{n=0}^{\infty} \frac{\lambda_n^2}{\lambda_1^2} |A_n|^2 e^{-2t\eta\lambda_n^2/a^2} \leq -2\eta\lambda_1^2 \epsilon \quad (18)$$

The Gronwall inequality leads finally to a minimal rate of variation for the energy:

$$\epsilon(t) \leq \epsilon(0) e^{\frac{2\eta\lambda_1^2}{a^2} t} \quad (19)$$

3 Dissipation in a cylindrical conductor-2

In the previous problem, the initial current lead to very simple boundary conditions preventing the magnetic field to come out from the cylinder. It is thus interesting to study a similar problem but with different boundary conditions.

In this section, we consider a problem with the same geometrical configuration than in the previous section but with a different initial current. At $t = 0$ (and $t = 0$ only), we impose an axial current depending only on r :

$$\mathbf{J}(\mathbf{r}) = J_z(r) \mathbf{e}_z \quad (20)$$

Obviously, the magnetic field created by this current will be of the form $B_\theta(t, r) \mathbf{e}_\theta$. Outside the conductor, the magnetic field is now non-zero and satisfy the equation $\nabla \times \mathbf{B} = 0$, yielding an harmonic field going to zero at infinity. Inside the conductor the field obey the diffusion equation:

$$\frac{\partial B_\theta}{\partial t} = \eta \left[-\frac{\partial^2 B_\theta}{\partial r^2} - \frac{1}{r} \frac{\partial B_\theta}{\partial r} + \frac{B_\theta}{r^2} \right] \quad (21)$$

As before, the decaying magnetic field is $B_\theta(t, r) = b(r)e^{-\alpha t}$. By using the change of variable $\rho = r\sqrt{\alpha/\eta}$, the equation (21) become:

$$\frac{\partial^2 b(\rho)}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial b(\rho)}{\partial \rho} + b(\rho) \cdot \left(1 - \frac{1}{\rho^2}\right) = 0 \quad (22)$$

This is the bessel equation for degree $m = 1$. A solution of the equation is thus:

$$b = J_1\left(\sqrt{\frac{\alpha}{\eta}}r\right) \quad (23)$$

In this case, the boundary conditions do not provide sufficient constraint on the magnetic field and thus do not lead to vanishing conditions for Bessel functions. This means that the system can not be represented by a Fourier-Bessel serie. In particular, the decay rates α can not be discretized, which seems to be in agreement with the fact that the energy is infinite outside the cylinder.

Using Hankel transform, we can however expand the inner magnetic field as follow :

$$b_i(r, t) = \int_0^\infty d\alpha c(\alpha) J_1\left(\sqrt{\frac{\alpha}{\eta}}r\right) e^{-\alpha t} \quad (24)$$

Using the transform $\beta = \sqrt{\alpha/\eta}$, the equation (24) becomes:

$$b_i(r, t) = 2\eta \int_0^\infty \beta d\beta C(\beta) J_1(\beta r) e^{-\eta\beta^2 t} \quad (25)$$

defined only for $r < a$. Outside the conductor, the magnetic field is simply harmonic and thus have the form $b_o(r, t) = A(t)/r$. The continuity of the tangential component of $H = B/\mu$ across the boundary yields an expression for the outer field:

$$b_o(r, t) = 2\eta \frac{\mu_0 a}{\mu r} \int_0^\infty \beta d\beta C(\beta) J_1(\beta a) e^{-\eta\beta^2 t} \quad (26)$$

We will now suppose that it is possible to find some magnetic field b_T defined on all the domain ($0 < r < \infty$) and represented on the basis of Bessel function:

$$b_T(r, t) = 2\eta \int_0^\infty \beta d\beta A(\beta) J_1(\beta r) e^{-\eta\beta^2 t} \quad (27)$$

In order to find an expression for the coefficient $A(\beta)$, we will evaluate the integral:

$$Y = \int_0^\infty r dr J_1(\beta r) b_T(r, 0) \quad (28)$$

There is two way to evaluate Y . By considering the expansion of B_T :

$$Y = \int_0^\infty r dr J_1(\beta r) 2\eta \int_0^\infty \beta' d\beta' A(\beta') J_1(\beta' r) e^{-\eta\beta'^2 t} \quad (29)$$

Using the orthogonality condition :

$$\int_0^\infty r dr J_1(\beta r) J_1(\beta' r) = \frac{\delta(\beta - \beta')}{\beta} \quad (30)$$

the equation (29) become:

$$Y = 2\eta A(\beta) \quad (31)$$

We can also evaluate Y by separating the domain of integration:

$$Y = \int_0^a r dr J_1(\beta r) b_i(r, 0) + \int_a^\infty r dr J_1(\beta r) b_o(r, 0) \quad (32)$$

$b_o(r, 0)$ is related to $b_i(r, 0)$ and we get:

$$Y = \int_0^a r dr J_1(\beta r) b_i(r, 0) + b_i(a, 0) \frac{\mu_0 a}{\mu} \int_a^\infty dr J_1(\beta r) \quad (33)$$

By grouping equations (31) and (33) we finally get an expression for the coefficient $A(\beta)$ of the magnetic field:

$$A(\beta) = \frac{1}{2\eta} \int_0^a r dr J_1(\beta r) b_i(r, 0) + b_i(a, 0) \frac{\mu_0 a}{2\eta\mu} \int_a^\infty dr J_1(\beta r) \quad (34)$$

By comparison with the first problem where decay modes were quantified, it can be interesting to investigate the evolution of the energy of the system. The energy in the infinite volume is given by:

$$\epsilon = \int_0^\infty r dr b_T^2 = \int_0^\infty \beta d\beta |A(\beta) e^{-\eta\beta^2 t}|^2 \quad (35)$$

We separate this integral in two parts, using a small parameter w :

$$\epsilon = \int_0^w \beta d\beta |A(\beta) e^{-\eta\beta^2 t}|^2 + \int_w^\infty \beta d\beta |A(\beta) e^{-\eta\beta^2 t}|^2 \quad (36)$$

In the first part of the integral, $A(\beta)$ can be expressed by its Taylor expansion $A(\beta) = K\beta$ since $A(0) = 0$. In the second part, the asymptotic limit of long time t allow us to neglect this term to zero. The energy is then given by:

$$\epsilon \sim \int_0^\infty \beta^3 d\beta e^{-2\eta\beta^2 t} \quad (37)$$

By integration by parts we get $\epsilon \sim t^{-3/2}$. We can easily derive this equation and we finally get the result :

$$\frac{\dot{\epsilon}}{\epsilon} \sim \frac{\mu}{t} \quad (38)$$

In the previous problem, the magnetic field was restricted to a finite radius, yielding a quantification of the decay modes. As a consequence, the energy was bounded by the exponential decay of the less damped mode. In this new problem, the magnetic field is no longer restricted to a finite volume but goes to infinity. Moreover, the total energy is infinite outside the conductor. We see here that it is thus impossible to get some exponential decay for the magnetic energy. In this situation, the system is damping the energy with a much slower rate. This is probably reminiscent from the infinite amount of energy to dissipate coming from outside and collapsing on the conductor.

4 Decay problem in spherical geometry

We have seen how the magnetic field is decaying in a cylindrical conductor under the action of ohmic diffusion. In some case, it is impossible to observe exponential bound of the magnetic energy. In this section we study the problem of free decay modes in a spherical configuration.

Let's consider a spherical conductor of permeability μ and conductivity σ surrounded by infinite vacuum. In such a geometry, it is in general useful to decompose the solenoidal fields into poloidal part P and toroidal part T . The magnetic field is given by :

$$\mathbf{B} = \nabla \times \nabla \times (\mathbf{r}P) + \nabla \times (\mathbf{r}T) \quad (39)$$

The governing equations for the toroidal part are:

$$\frac{\partial T}{\partial t} = \eta \Delta T \quad r < a \quad (40)$$

$$T = 0 \quad r > a \quad (41)$$

Similarly we get two equations for the poloidal part:

$$\frac{\partial P}{\partial t} = \eta \Delta P \quad r < a \quad (42)$$

$$\Delta P = 0 \quad r > a \quad (43)$$

This decomposition in toroidal and poloidal part is interesting because P and T may always be decomposed on the basis of the spherical harmonics, which are the eigenfunction of the angular part of the spherical laplace operator. The problem for T and P are totally independent and can be studied separately. I will present all the work as axisymmetric problem for simplicity but the results are totally similar for non-axisymmetric problem.

4.1 Toroidal decay problem

The scalar T can be decomposed as follow:

$$T = \sum_l \hat{T}^l Y_l(\theta, \phi) \quad (44)$$

Using this decomposition the equation (40) become:

$$\frac{\partial \hat{T}^l}{\partial t} = \eta \left(\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \hat{T}^l}{\partial r} \right) - \frac{l(l+1)}{r^2} \hat{T}^l \right) \quad (45)$$

We know that it is a decaying problem so we can suppose $\hat{T}^l(r, t) = T^l(r) e^{-\alpha t}$. We then get the equation:

$$\frac{\partial^2 T^l}{\partial r^2} + \frac{2}{r} \frac{\partial T^l}{\partial r} + \left(\frac{\alpha}{\eta} + \frac{l(l+1)}{r^2} \right) T^l = 0 \quad (46)$$

This equation is the spherical bessel equation and a solution to this equation is :

$$T_\alpha^l = r^{-1/2} J_{l+\frac{1}{2}} \left(\sqrt{\frac{\alpha}{\eta}} r \right) \quad (47)$$

Boundary conditions for toroidal magnetic field are:

$$T^l(a) = 0 \implies J_{l+\frac{1}{2}} \left(\sqrt{\frac{\alpha_n}{\eta}} a \right) = 0 \quad (48)$$

As for the problem 1, the vanishing of T at the boundary yield a determination of the zeros $\lambda_n = \sqrt{\frac{\alpha_n}{\eta}} a$ of the Bessel function. We can then expand T^l as a Fourier-Bessel serie:

$$T^l = \sum_{n=0}^{\infty} A_n r^{-1/2} e^{-t\eta\lambda_n^2/a^2} J_{l+\frac{1}{2}}(\lambda_n r) \quad (49)$$

The total representation of the toroidal magnetic field inside the conductor is thus:

$$T_i(r, t) = \sum_l Y_l(\theta, \phi) \sum_{n=0}^{\infty} A_n^l r^{-1/2} e^{-t\eta(\lambda_n^l)^2/a^2} J_{l+\frac{1}{2}}(\lambda_n^l r) \quad (50)$$

The field is now expanded in term of Bessel function but also of spherical harmonics. In addition of the orthogonality for bessel function (9) we have to consider the orthogonality of spherical harmonic:

$$\int d\Omega Y_l^m(\theta, \phi) Y_{l'}^{m'}(\theta, \phi) = \delta_{ll'} \delta_{mm'} \quad (51)$$

We can now use boundary conditions to compute the coefficient A_n^l . Because the toroidal field is vanishing on the boundary, it leads to consider integral restricted to the conducting domain. We evaluate thus the integral :

$$I = \int d\Omega \int_0^1 r^2 dr T_i(r, 0) r^{-1/2} J_{l+\frac{1}{2}}(\lambda_n^l r) Y_l(\theta, \phi) \quad (52)$$

We expanding $T_i(r, 0)$ and using orthogonality, I becomes:

$$I = \int d\Omega \int_0^1 r^2 dr \sum_{l'} \sum_{n'} A_{n'}^{l'} (r^{-1/2})^2 J_{l+\frac{1}{2}}(\lambda_n^l r) J_{l'+\frac{1}{2}}(\lambda_{n'}^{l'} r) Y_l(\theta, \phi) Y_{l'}(\theta', \phi) \quad (53)$$

$$I = \sum_{l'} \int d\Omega Y_l(\theta, \phi) Y_{l'}(\theta', \phi) \sum_{n'} A_{n'}^{l'} \int_0^1 r dr J_{l+\frac{1}{2}}(\lambda_n^l r) J_{l'+\frac{1}{2}}(\lambda_{n'}^{l'} r) \quad (54)$$

$$I = A_n^l [J_{l+\frac{3}{2}}(\lambda_n^l)]^2 \quad (55)$$

By changing bounds of the integral over r, we finally obtain an expression for A_n^l :

$$A_n^l = \frac{\sqrt{a\eta}}{a^3 [J_{l+\frac{3}{2}}(\lambda_n^l)]^2} \int dV_{cond} r^{-1/2} J_{l+\frac{1}{2}}(\lambda_n^l r/a) \quad (56)$$

We see here that the boundary conditions for the toroidal field lead to quantification of the modes and, as in the section 2, we expect that the energy will be dominated by the decay rate of the modes. Using orthogonality of Bessel fuction and spherical harmonic we get an expression for the evolution of the energy:

$$\epsilon_T = \sum_l \sum_n \frac{|A_n^l|^2}{2} e^{-2\eta(\lambda_n^l)^2 t/a^2} [J_{l+\frac{3}{2}}(\lambda_n^l)]^2 \quad (57)$$

By derivating this energy with respect to t and using Gronwall inequality like in section 2, we get an exponential bounds for the energy related to the less damped mode:

$$\epsilon_T \leq \epsilon_T(0) e^{-2\eta(\lambda_1^1)^2/a^2} \quad (58)$$

4.2 Poloidal decay problem

Let's do the same thing for the poloidal part of the magnetic field. The SPherical harmonic decomposition is still possible:

$$P = \sum_l \hat{P}^l Y_l(\theta, \phi) \quad (59)$$

Inside the conductor, the equation is the same than for the toroidal part and we get the solution:

$$P_\gamma^l = r^{-1/2} J_{l+\frac{1}{2}}(\sqrt{\frac{\gamma}{\eta}} r) \quad (60)$$

Where γ denotes now the decay rate of poloidal component P_γ^l . The boundary conditions are however different in the case of the poloidal field. According to the equation (43) the scalar P is harmonic outside the conductor and $P^l(r, t) = c(t)r^{-(l+1)}$. The radial derivative of this field at the outer boundary is then related to P by:

$$\frac{\partial P_\gamma^l}{\partial r} + \frac{(l+1)}{a} P = 0 \quad (61)$$

By plugging expression (60) into equation (61) and using recurrence relation for the Bessel functions, we obtain:

$$J_{l-\frac{1}{2}}(\sqrt{\frac{\gamma}{\eta}} a) = 0 \quad (62)$$

We see here that the boundary conditions discretize the possible decay rate of the field as for toroidal part but involve now zero of $J_{l-\frac{1}{2}}$. This is easy to relate it to the zeros λ_n^l of $J_{l-\frac{1}{2}}$ and the decay rates of the modes are given by:

$$\gamma_n^l = \eta \frac{(\lambda_n^{l-1})^2}{a^2} \quad (63)$$

We thus have the following expansion for P inside the conductor:

$$P_i(r, t) = \sum_l Y_l(\theta, \phi) \sum_{n=0}^{\infty} B_n^l r^{-1/2} e^{-t\eta(\lambda_n^{l-1})^2/a^2} J_{l+\frac{1}{2}}(\lambda_n^{l-1} r) \quad (64)$$

Outside the conductor, the field is expanding only on the spherical harmonics, and is totally determined by the value of the internal field at the boundary:

$$P_o(r, t) = \sum_l c^l(t) Y_l(\theta, \phi) r^{-(l+1)} \quad (65)$$

Where we use the boundary conditions to calculate the coefficient $c^l(t)$:

$$c^l(t) = \sum_{n=0}^{\infty} B_n^l a^{l+1/2} e^{-\eta(\lambda_n^{l-1})^2/a^2} J_{l+\frac{1}{2}}(\lambda_n^{l-1}a) \quad (66)$$

The determination of the coefficient is identical as for the toroidal field and give:

$$B_n^l = \frac{\sqrt{a\eta}}{a^3 [J_{l+\frac{3}{2}}(\lambda_n^{l-1})]^2} \int dV_{cond} r^{-1/2} J_{l+\frac{1}{2}}(\lambda_n^{l-1}r/a) \quad (67)$$

Here again, the rate of decay of the magnetic energy will be governed by the decay rate of the less damped mode. However, because of the shift in the index l , the first eigenvalue is now given by $\lambda_1^0 = \pi$ and we get the following bounds for the rate of decay of the energy:

$$\epsilon_P \leq \epsilon_P(0) e^{-\frac{2\eta\pi^2}{a^2}t} \quad (68)$$

Although the poloidal magnetic field is not constrained to a finite volume like the toroidal one, we see that it can also be represented as a discrete sum of magnetic modes, with quantification of the possible decay rate, yielding an exponential bound for the energy. We note however that the amount of energy outside is finite, in opposition with the problem of the section 3 where an exponential decay could not be reached.

By comparing equations (58) and (68), we note that the poloidal energy is decaying much slowly than the toroidal field. This can easily be explained: P is not vanishing outside and create a large amount of energy. In the vacuum, there is no mechanisms to dissipate the energy and this energy is forced to collapse on the conductor by use of the Poynting flux. In consequence it is more difficult for the system to dissipate this energy. In the limit of an infinite energy, we have seen in section 3 that the system lost its ability to exponentially damp the energy and we get a power law.

The conclusions presented here simply a conjecture of the different case explored here and can not be taken as rigorous assumption. We will next use a more simple model trying to capture the essential arguments presented here.

5 Simple model

The geometries and the calculations involved in the previous cases are relatively complicated. In this section we will study a very idealized mathematical model, in the perspective of a better comprehension of the different behavior of the energy depending on the situation outside the conductor.

Let us consider a one dimensional problem involving a scalar $\phi(x, t)$ depending only on the x direction. The domain is divided in two part: the inside part for $r < a$ and the outside part from a to infinity. Inside, ϕ obey the equation:

$$\dot{\phi} = D\phi'' \quad (69)$$

Where the dot means time derivative and the prime means x derivative. D is a coefficient of diffusion. Outside the conducting region, we suppose an equation of the type:

$$\phi(x)'' - m^2\phi = 0 \quad (70)$$

Where m is a parameter representing the spatial damping of the field outside. Indeed solution outside is of the form $\phi \sim e^{-m(x-a)}$. By supposing exponential decay for the time dependance of ϕ we get the equation :

$$\phi(x)'' + \frac{\alpha}{D}\phi = 0 \quad (71)$$

α is the decay rate of the mode. The boundary condition for the field at the origin is $\phi'(0) = 0$. A solution of this equation satisfying ϕ non-zero at $x = 0$ is $\phi(x) = \cos(kx)$ with $k = \sqrt{\alpha/D}$. We can now use the boundary conditions for this problem. We suppose here that ϕ and its derivative in x are continuous at the interface, leading to the same kind of relation than in the case of the poloidal field:

$$\phi' + m\phi = 0 \quad (72)$$

By plugging the expression for ϕ in this equation, we get:

$$\tan(k_na) = \frac{m}{k_n} \quad (73)$$

This yield a quantification of the decay rate $\alpha_n = k_n^2 D$. We can now expand our field on the Fourier-cosine series:

$$\phi(x, t) = \sum_n \phi_n \cos(k_n x) e^{-k_n^2 D t} \quad (74)$$

By using of the classical orthogonality relationship for cosine, the coefficients are obviously given by:

$$\phi_n = \int_0^a dx \cos(k_n x) \phi(x, 0) \quad (75)$$

As in the previous section, the energy is bounded by the less damped mode and we have:

$$\epsilon_\phi \leq \epsilon_\phi(0) e^{-2Dk_1^2 t} \quad (76)$$

This model show clearly how the decay rate of the energy of a field is related to the extension of this field in the space. Here, m represent the damping of the field outside the conductor. For large m , the field is strongly damped near the boundary and tend to a constant value of $(\frac{\pi}{2a})^2$ when m is increased. This is the situation for the diffusion of toroidal field seen before, where the field is confined to the inside region. When m is decrease, this correspond to an increase of the extension of the field outside the conductive region. The slowest decay

rate is then reducing and the corresponding energy is less and less damped. In the limit of m very small, the decay rate of the energy tends to zero. However, one can note that when $m = 0$, the only solution satisfying the boundary conditions is $\phi = 0$. This behavior illustrates all the case studied before and show clearly the relationship between the external energy and the internal diffusion.

Using these simple examples, we have seen how the energy outside a conductor can play an essential role in the dissipation of the energy inside the conductor. In particular, a large amount of energy outside the conductor will in general collapse on the conductor and reduce the damping of the total energy by forcing the system to dissipate more energy. This means that surrounding a conductor with vacuum can help the system to keep its energy. In experimental situation, for instance in dynamo experiment, the conducting fluid is in general separated from the insulating region by the container, which can be a metal with different magnetic properties. In the next section we will study the decay problem for this more complicated situation.

6 Effect of ferromagnetic materials

We consider the same situation than for the decay problem in spherical geometry. In addition, we suppose now that there is two concentric sphere with region of different permeability. We consider only the decay problem for a poloidal magnetic field. The equation in the inner sphere of radius $r = r_1$ is given by:

$$\frac{\partial P}{\partial t} = \eta_1 \Delta P \quad (77)$$

For the external shell ($r_1 < r < r_2$) we have the equation :

$$\frac{\partial P}{\partial t} = \eta_1 \Delta P \quad r < a \quad (78)$$

For $r > r_2$, the field is harmonic in a medium of permeability μ_0 . By using spherical harmonic expansion and solving the radial equation we find that the solution to this problem, for a given decay rate α , is given by:

$$P^l(r) = Cr^{-1/2} J_{l+\frac{1}{2}} \left(\sqrt{\frac{\alpha}{\eta_1}} r \right) \quad (79)$$

for the field in the inner sphere. In the shell we have now:

$$P(r) = Ar^{-1/2} J_{l+\frac{1}{2}} \left(\sqrt{\frac{\alpha}{\eta_2}} r \right) + Br^{-1/2} N_{l+\frac{1}{2}} \left(\sqrt{\frac{\alpha}{\eta_2}} r \right) \quad (80)$$

where $N_l(r)$ is the neumann function which appears because we can not invoke singularity at the center for eliminating it. Note that the decay rate is assumed to be identical in the two region of the problem. The boundary conditions at the external sphere is the same than before, but we suppose now that the permability of the two regions are not identical:

$$\frac{\partial P}{\partial r} + \frac{\mu_2}{\mu_0} \frac{(l+1)}{r_2} P = 0 \quad (81)$$

Using Bessel expansion it gives us the relation:

$$kr_2 J_{l-\frac{1}{2}}(kr_2) + (l+1)\left(\frac{\mu_2}{\mu_0} - 1\right) J_{l+\frac{1}{2}}(kr_2) = -\frac{B}{A} \left[kr_2 N_{l-\frac{1}{2}}(kr_2) + (l+1)\left(\frac{\mu_2}{\mu_0} - 1\right) N_{l+\frac{1}{2}}(kr_2) \right] \quad (82)$$

with $k = \sqrt{\frac{\alpha}{\eta_2}}$ The second boundary condition, acting on the inner core, consists on the following system:

$$[B] = 0 \quad (83)$$

$$\left[\frac{\partial B}{\partial r} \right] = \frac{\mu_1}{\mu_2} \quad (84)$$

leading to the equation :

$$\frac{B}{A} = \frac{\sqrt{\frac{\mu_2 \sigma_1}{\mu_1 \sigma_2}} j_l(\rho_2) j'_l(\rho_1) - j'_l(\rho_2) j_l(\rho_1)}{-\sqrt{\frac{\mu_2 \sigma_1}{\mu_1 \sigma_2}} y_l(\rho_2) j'_l(\rho_1) - j'_l(\rho_2) y_l(\rho_1)} \quad (85)$$

The decay rates are then quantized and given by solving the system of equations (82) and (85). This can be done numerically for any value of the permeabilities. However, by considering the limit of high permeability for the shell, asymptotic calculation leads to the following expression for the decay rates of the system:

$$\alpha_l^n = (l+n)^2 \frac{\pi^2 \sigma_1 \mu_1}{(r_1 + r_2)^2 \sigma_2 \mu_2} \quad (86)$$

We see here that the limit of high permeability leads to a decrease of the damping rate of the energy of the system. This effect is stronger for a thin shell. This result suggests interesting modification for dynamo experiment in spherical geometry currently ongoing. For example, in the Madison experiment, surrounding the experiment with a shell of ferromagnetic materials (soft iron for instance) will probably yield a more excitable system and may cause a decreasing in the threshold of the dynamo instability (see figure ??). Increasing the permeability of the inner sphere instead of the external ones leads to similar decrease of the decay rate and using ferromagnetic core would be a modification for dynamos experiment like DTS experiment in Grenoble in France or in Maryland experiment where two concentric spheres are used (see figure ??).

In more realistic MHD situation, the conductor is a fluid, the system has now more complex way to dissipate the energy. It is thus of primary interest to consider the MHD case and see how the system dissipates the energy.

7 MHD bounds and background method

In this section we would like to study shear flow of a conducting fluid across a magnetic field. This situation occurs in many physical situations in astrophysical objects or industrial applications. In general, this implies the existence of Hartmann layers near the boundaries of the system. In particular we want to study such a Hartmann layer when one considers

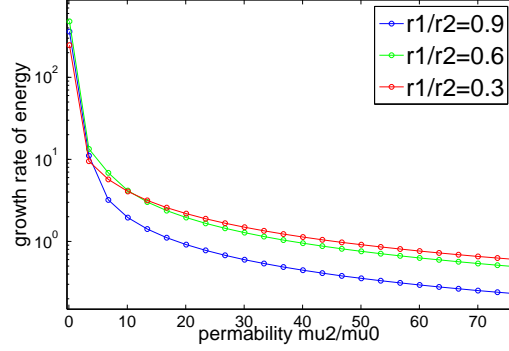


Figure 1: decay rate of the energy inside the sphere when the permeability of the shell is increased, for different aspect ratio. In the limit of ferromagnetic shell, the energy of the system do not decay.

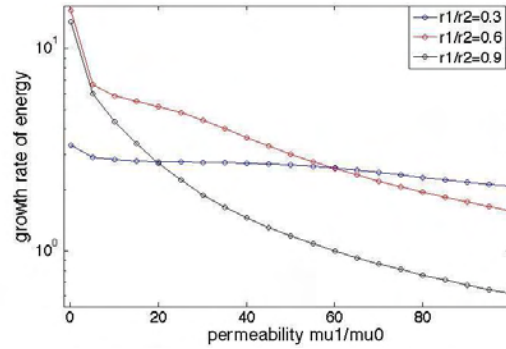


Figure 2: decay rate of the energy inside the sphere when the permeability of the core is increased, for different aspect ratio. In the limit of ferromagnetic core, the energy of the system do not decay.

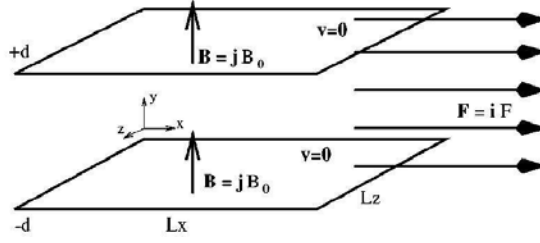


Figure 3: Setup of the Hartmann problem

realistic boundary conditions for the magnetic field, for instance when the conducting region is in contact with insulator walls or vacuum. The main quantities of interest are the energy stability and the rate of dissipation of this system.

The problem consists of a layer of conducting fluid of kinematic viscosity ν and magnetic resistivity η confined between two horizontal plates of size L in x and z directions. The plates are located between $y = -d/2$ and $y = +d/2$. There is a force F creating some flow in the x direction and an imposed constant magnetic field B_0 in the vertical direction. We assume periodic boundary conditions in the horizontal directions x and z for all the fields. In the vertical direction y there is no slip boundary conditions for the velocity field and the magnetic field match a potential field due to insulating region outside (see figure ??).

The solenoidal velocity fields \mathbf{v} and \mathbf{B} satisfy the MHD equations:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla \pi + \nu \Delta \mathbf{v} + \mathbf{F} + \frac{1}{\mu \rho} (\mathbf{B} \cdot \nabla) \mathbf{B}, \quad (87)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \Delta \mathbf{B}. \quad (88)$$

In the above equations, ρ is the density, μ is the magnetic permeability and σ is the conductivity of the fluid.

A laminar solution for Hartmann flow is given by:

$$U_H = \frac{Fd}{2B_0} \sqrt{\frac{\eta}{\nu}} \left[\frac{\cosh(\frac{B_0 d}{2\sqrt{\nu\eta}}) - \cosh(\frac{B_0 y}{\sqrt{\nu\eta}})}{\sinh(\frac{B_0 d}{2\sqrt{\nu\eta}})} \right] \quad (89)$$

for the velocity field and :

$$B_H = \frac{Fd}{2B_0} \left[\frac{\sinh(\frac{B_0 y}{\sqrt{\nu\eta}})}{\sinh(\frac{B_0 d}{2\sqrt{\nu\eta}})} - \frac{2y}{d} \right] \quad (90)$$

For the magnetic field.

We will use different dimensionless number in this section. The Hartmann number:

$$Ha = \frac{B_0 d}{\sqrt{\nu\eta}} \quad (91)$$

The Grashoff number:

$$Gr = \frac{Fd^3}{\nu^2} \quad (92)$$

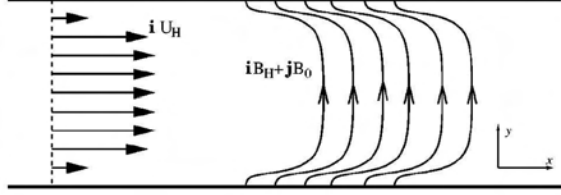


Figure 4: Laminar solution for the Hartmann problem

and the magnetic grashoff number:

$$G_m = \frac{F d^3}{\nu \eta} \quad (93)$$

We define the dissipation associated with this laminar state as:

$$D = \nu \langle |\nabla \mathbf{U}_H|^2 \rangle + \eta \langle |\nabla \mathbf{B}_H|^2 \rangle \quad (94)$$

where brackets indicate space averaging. We can then evaluate the laminar dissipation:

$$D = \frac{H a^{1/2}}{2\sqrt{2}H a} \left[\coth(H a) - \frac{1}{H a} \right] \quad (95)$$

We note here that the magnetic field tends to zero when the applied field vanish and the velocity reduce to Poiseuille flow.

7.1 Background decomposition

In order to do the energy stability and estimate bounds on the dissipation, an useful method is to decompose the fields into two parts. A stationnary background profile depending only on y and a perturbation part:

$$\mathbf{v} = \mathbf{i} U_b(y) + \mathbf{u} \quad (96)$$

$$\mathbf{B} = \mathbf{i} B_b(y) + \mathbf{j} B_o + \mathbf{b} \quad (97)$$

The background profile and the perturbations can be any fields given the conditions that the total field have to be divergence-free and satisfy the boundary conditions. We can then use this decomposition in the equations (87) and (88). By adding the two equations, this gives the expression for the energy of the perturbation defined by $E = \mathbf{u}^2/2 + \mathbf{b}^2/2$:

$$\begin{aligned} \frac{\partial E}{\partial t} = & \langle u_x B_0 B'_b \rangle + \langle b_x B_0 U'_b \rangle + \langle (b_x b_y - u_x u_y) U'_b \rangle \\ & + \langle (u_x b_y - u_y b_x) B'_b \rangle + \eta \langle \Delta \mathbf{b} \cdot \mathbf{b} \rangle - \nu \langle |\nabla \mathbf{u}|^2 \rangle \\ & + \nu \langle u_x U''_b \rangle + \eta \langle b_x B''_b \rangle + \langle u_x F \rangle \end{aligned} \quad (98)$$

7.2 Poincare inequality

Before doing the energy stability, we need to obtain an inequality for the term $\langle \Delta \mathbf{b} \cdot \mathbf{b} \rangle$ in the energy equation above. Indeed, the classical Poincare inequality which we use for the

velocity field can not apply for the magnetic field due to the boundary conditions. Let's derive the magnetic field in Fourier space:

$$b_i = \sum_k b_i^k(y) \phi_k \quad (99)$$

i.e.

$$b_i = \sum_{k=0}^{\infty} b_i^k(y) e^{i(k_x x + k_z z)} \quad (100)$$

where subscript i stands for x, y or z and ϕ_k are the eigenfunction of the horizontal laplacian with eigenvalue $\lambda_k = k_x^2 + k_z^2 = k_h^2$, satisfying the equation:

$$(\partial_{xx} + \partial_{zz})\phi_k = \lambda_k \phi_k \quad (101)$$

We see that the fourier coefficient depend on y since b is different from zero at the boundaries and Fourier decomposition is impossible in this direction. However, the field is harmonic outside the conductor and the Fourier modes satisfy :

$$(\partial_{yy} - k_h^2)b_i^k(y) = 0 \quad (102)$$

which give us :

$$b_i^k(y) = e^{-k_h(y-d/2)} \quad (103)$$

$$b_i^k(y) = e^{k_h(y+d/2)} \quad (104)$$

In the same manner that for the poloidal field in the sphere, the continuity of the field and its derivative imply :

$$\partial_y b_i^k\left(\frac{d}{2}\right) + k_h b_i^k\left(\frac{d}{2}\right) = 0 \quad (105)$$

$$\partial_y b_i^k\left(-\frac{d}{2}\right) - k_h b_i^k\left(-\frac{d}{2}\right) = 0 \quad (106)$$

We remark here that the magnetic field satisfy mode by mode exactly the same conditions than the toy model. Suppose now that the vertical dependance of the field can be described by the following expansion for the coefficients $b_i^k(y)$:

$$b_i^k(y) = A \cos(q_k y) + B \sin(q_k y) \quad (107)$$

The boundary conditions becomes:

$$\tan\left(\frac{q_k d}{2}\right) = \pm \left(\frac{k_h}{q_k}\right)^{\pm 1} \quad (108)$$

This equation yield an expression for the wavenumber q_k . Using this expansion, it is straightforward to derive an expression for $\Delta B \cdot B$:

$$\begin{aligned} -\langle \Delta B \cdot B \rangle &= - \int_{-\frac{d}{2}}^{+\frac{d}{2}} dy \sum_k b^k (\partial_{yy} - k_h^2) b^k \\ &= \int_{-\frac{d}{2}}^{+\frac{d}{2}} dy \sum_k (q_k^2 + k_h^2) |b^k|^2 \geq (q_1^2 + k_1^2) \langle B^2 \rangle \end{aligned} \quad (109)$$

Where we use the first wavenumber to bound the expression. Eliminate the $k = 0$ terms, (there is no horizontally averaged magnetic field) we get the Poincare inequality :

$$-\eta \langle \Delta \mathbf{B} \cdot \mathbf{B} \rangle \geq \eta \left(\frac{\pi^2}{L^2} + q_1^2 \right) \langle B^2 \rangle \quad (110)$$

7.3 Energy stability

The stability of the Hartmann layer is an important characteristic of the system. A complete determination of the problem would imply to solve the full variationnal problem involving the resolution of the Euler-Lagrange equation. We will instead derive rigourous energy stability by considering appropriate bounds on the energy. Note that this method generally capture the main behavior of the system and the full resolution only improve the results by some factor.

The background kinetic and magnetic fields are now taken to be the laminar solutions of the problem. The energy equation (98) simplify in :

$$\frac{\partial E}{\partial t} = -F[\mathbf{u}, \mathbf{b}] \quad (111)$$

where the funtional F is given by:

$$\begin{aligned} F[\mathbf{u}, \mathbf{b}] = & \langle (b_x b_y - u_x u_y) U'_H \rangle + \langle (u_x b_y - u_y b_x) B'_H \rangle \\ & - \eta \langle \Delta \mathbf{b} \cdot \mathbf{b} \rangle + \nu \langle |\nabla \mathbf{u}|^2 \rangle \end{aligned} \quad (112)$$

Using Holster and Young inequality, we can bound the two first term of the above quaratic form as follow:

$$\langle (b_x b_y - u_x u_y) U'_H \rangle \leq \frac{\max(|U'_H|)}{2} \langle \mathbf{b}^2 + \mathbf{u}^2 \rangle \quad (113)$$

$$\langle (u_x b_y - u_y b_x) B'_H \rangle \leq \frac{\max(|B'_H|)}{2} \langle \mathbf{b}^2 + \mathbf{u}^2 \rangle \quad (114)$$

The two last term can be controlled by Poincare inequalities

$$\nu \langle |\nabla \mathbf{u}|^2 \rangle \geq \nu \frac{\pi^2}{d^2} \langle \mathbf{u}^2 \rangle \quad (115)$$

$$-\eta \langle \Delta \mathbf{b} \cdot \mathbf{b} \rangle \geq \eta \left(\frac{\pi^2}{L^2} + q_0^2 \right) \langle \mathbf{b}^2 \rangle \quad (116)$$

The above inequality yield conditions for non-negativity of the quadratic form (112) which can be formulated as:

$$G_m^2 [Ha \coth(Ha) - 1]^2 \leq Ha^2 (2\pi^2 - G_r) \left(2 \left(\frac{d^2 \pi^2}{L^2} + q_0^2 d^2 \right) - G_m \right) \quad (117)$$

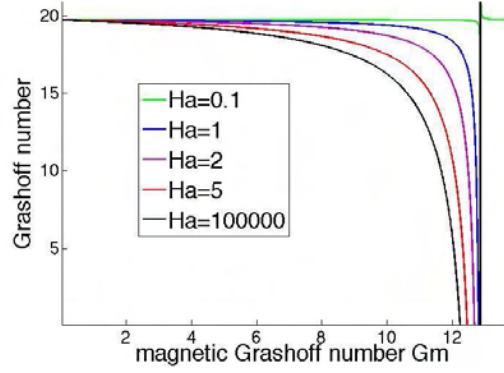


Figure 5: Energy stability of the Hartmann flow for different Hartmann number.

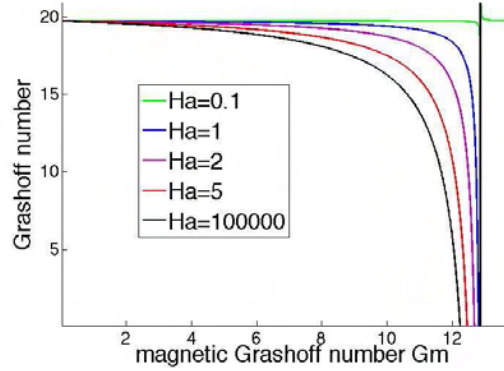


Figure 6: Evolution of the stability with the aspect ratio of the problem.

In the limit where G_m tends to zero, we find the energy stability of the Poiseuille flow. Figure (??) show the energy stability in the plane $G_r - G_m$ for different Hartmann number. We see here that the results depend now on the aspect ratio $r = L/d$ and the figure (??) illustrate the destabilisation of the Hartmann flow when the length perpendicular to the magnetic field is increased.

7.4 Bound on dissipation rate

Using the procedure followed in the previous section, we are now interesting in the total dissipation of the system, defined by:

$$D = \nu \langle |\nabla \mathbf{v}|^2 \rangle + \eta \langle |\nabla \mathbf{B}|^2 \rangle \quad (118)$$

To determined bounds on the dissipation, we now choose arbitrary backgrounds, which are not the laminar solutions but which will be determined in order to get the best possible estimation of the bound.

Taking the equation (98) and adding to the energy half of the total dissipation we obtain:

$$\begin{aligned}
\frac{\partial E}{\partial t} + \frac{D}{2} = & \langle u_x B_0 B'_b \rangle + \langle b_x B_0 U'_b \rangle + \langle (b_x b_y - u_x u_y) U'_b \rangle \\
& + \langle (u_x b_y - u_y b_x) B'_b \rangle - \frac{\eta}{2} \langle \Delta \mathbf{b} \cdot \mathbf{b} \rangle - \frac{\nu}{2} \langle |\nabla \mathbf{u}|^2 \rangle \\
& + \frac{\nu}{2} \langle U_b'^2 \rangle + \frac{\eta}{2} \langle B_b'^2 \rangle + \langle u_x F \rangle \quad (119)
\end{aligned}$$

In order to eliminate the linear term in the previous expression, we perform the following change of variable:

$$\mathbf{u} = \mathbf{w} - \mathbf{i} V_b(y) \quad (120)$$

$$\mathbf{b} = \mathbf{h} - \mathbf{i} H_b(y) \quad (121)$$

with:

$$\nu V_b'' = B_0 B'_h \quad (122)$$

and:

$$\eta H_b'' = B_0 U'_h \quad (123)$$

We end up with the final expression for the dissipation rate of the Hartmann layer:

$$D = 2F \langle U_b \rangle - D_b + Q_b \quad (124)$$

where the functional Q is given by

$$Q_b = \nu \langle |\nabla \mathbf{w}|^2 \rangle - \eta \langle \Delta \mathbf{h} \cdot \mathbf{h} \rangle - 2 \langle (h_x h_y - w_x w_y) U'_b + (w_x h_y - w_y h_x) B'_b \rangle \quad (125)$$

We see that if we can find background fields such that the quadratic form Q_b is non-negative, the dissipation will be bounded by an expression depending only on the background profiles.

Using the same inequalities than for the estimation of the energy stability, we can obtain conditions for positivity of the quadratic form Q_b leading to a lower bound for the dissipation. This bound is valid only in the region of parameter space defined by

$$Rm \leq 4 \left(\frac{\pi^2 d^2}{L^2} + d^2 q_0^2 \right) \frac{1}{Re} \quad (126)$$

One can note that this prediction on the dissipation rate of the system is very weak. Indeed, the bound on the dissipation is restricted to a laminar region and do not apply for turbulent behavior. This problem come from the fact that the estimation of the bounds is done without taking into account the presence of the boundary layers which would occur in any physical situation.

Focusing on the boundary of the system, we can use the fundamental theorem of calculus, Schwartz and Young inequalities in order to get:

$$\langle (w_x w_y) U'_b \rangle \leq U^* \delta |\nabla \mathbf{w}|^2 \quad (127)$$

The derivation of this equation come from the fact that the conditions of vanishing perturbation w on the boundary give a control on the magnitude of the gradient of the field

inside the conductor. The same derivation is not possible for the magnetic field, where the matching with an external potential field do not offer any control on the value of h at the boundary. In fact, these boundary conditions specify only a relation between the field and its gradient at the boundary, which is not sufficient to conclude something about the gradient inside the conductor. In consequence, it is not possible to characterize the behavior of the magnetic field in the boundary layers, leading to the impossibility of obtaining a bound on the dissipation rate.

8 Conclusion

We have seen in the first sections that the vacuum can play an important role in the dissipation of a conductor or MHD system. It is thus of primary interest to see how stability or dissipation rate are changed when one take into account realistic boundary conditions. Using the background we studied the Hartmann flow problem in order to derive energy stability and lower bound on the dissipation rate. We have seen that the insulating boundary conditions do not allow sufficient control on the magnetic field in the boundary layer to get a bound on the dissipation rate. This suggest to use a different approach, involving a modification of the classical background method.

MHD Toy Model of the Solar Radiative Zone

Céline Guervilly
advised by Pascale Garaud

January 12, 2009

1 Introduction

The Sun can be roughly divided into three different regions:

- The nuclear core, located in the inner part until 20% of the solar radius¹ where heat is produced by nuclear fusion of hydrogen.
- Between $0.2r_{\odot}$ and $0.7r_{\odot}$ the temperature is not high enough for nuclear reactions to proceed. The energy produced in the core is transported through this region by radiation which gives its name to this layer: the radiative zone. This layer is characterised by strong thermal stratification.
- In the outer region heat is transported primarily by convective motions. This leads to a well-mixed layer called the convective zone.

An important feature of the solar interior is its rotation profile. In this report, we describe a toy model of the solar radiative zone that can be used to investigate the interior rotation rate. During the last few decades, technological advances in helioseismology have provided accurate observations of the solar interior. The axisymmetric angular velocity profile deduced from observations is shown in figure 1. Differential rotation is observed within the convective zone: the equatorial region rotates about 30% faster than the polar regions. There, the rotation profile at a fixed radius can be fairly accurately expressed as

$$\Omega_{cz}(\theta) = \Omega_{eq}(1 - a \cos^2 \theta - b \cos^4 \theta), \quad (1)$$

where Ω_{eq} is the rotation rate at the equator and θ is the colatitude. The numbers a and b are determined by the observations. Near the lower region of the convective zone $a = 0.17$ and $b = 0.08$ (Schou *et al.* [4]). Meanwhile the radiative zone is near-uniformly rotating at a rate $\Omega_{rz} \simeq 0.93\Omega_{eq}$ where Ω_{eq} is the angular velocity at the equator at the bottom of the convective zone. The transition between the convective zone and the radiative zone is called the tachocline and has an important dynamical role.

An obvious question which comes to mind is why is the radiative zone rotating uniformly.

¹Hereafter the solar radius is denoted r_{\odot} .

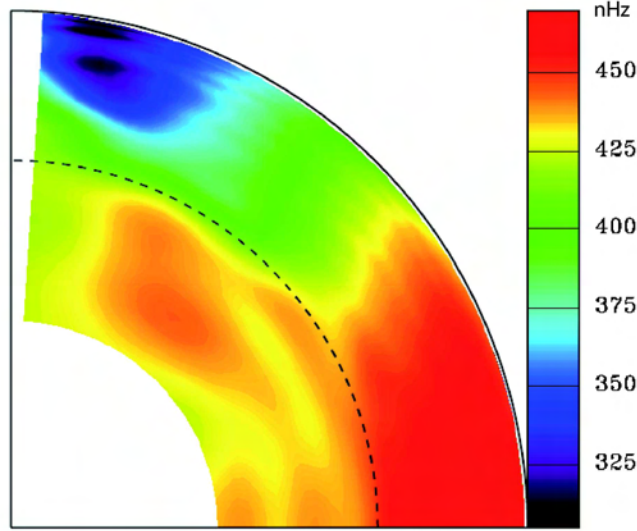


Figure 1: Angular velocity profile from solar observations in a meridional plane zonally and temporally averaged. The outer layer is the convective zone and the inner layer is the radiative zone. The dashed line represents the base of the convection zone. Observations of the inner most regions are not available by helioseismology. Source: SOHO/MDI

A standard explanation for this feature assumes the presence of a primordial magnetic field \mathbf{B} confined within the radiative zone. Ferraro’s isorotation law (1937), which is valid in the limit of negligible magnetic dissipation, states that

$$\mathbf{B} \cdot \nabla \Omega = 0, \quad (2)$$

for an axisymmetric rotating fluid in the steady state. Consequently the angular velocity is constant along magnetic field lines. If the field within the radiative zone has an “open” configuration, the magnetic field lines connected with the convective zone would transfer the differential rotation into the radiative zone; we shall call this a differentially rotating Ferraro state. On the otherhand, a close field configuration allows an uniformly rotating radiative zone. Gough and McIntyre [2] proposed a theoretical model (figure 2) that insured the uniform rotation of the radiative zone with:

- a differential rotation imposed by the convective zone,
- a primordial dipolar magnetic field in the radiative zone,
- down-welling meridional flows at the top of the radiative zone that confine the magnetic field, which itself prevents the flow from penetrating deeper into the radiative zone.

Numerical simulations based on this model have been carried out by Garaud & Garaud [1]. Their code is steady-state, axisymmetric, fully non-linear and models the bulk of the radiative zone as an anelastic conducting fluid. A radial velocity profile is imposed at the

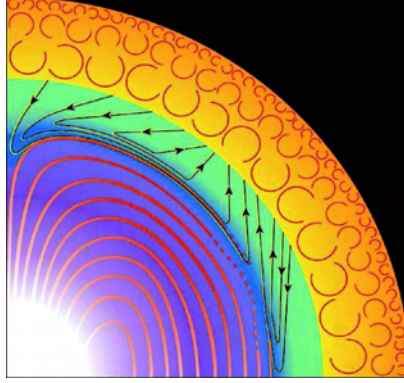


Figure 2: Picture of the model of Gough and McIntyre in a quadrant of the meridional plane. The red lines in the radiative zone (blue layer) represent the confined magnetic field, the black lines represent the meridional circulation which stays trapped in the top of the radiative zone. The yellow top layer is the convective zone.

top of the radiative zone and a no-slip condition is taken at the interface with the nuclear core. Both boundaries are assumed to be conducting with the same conductivity as the fluid. For low enough diffusivities they find that the rotation rate of the uniformly rotating region converges to about 86% of the imposed boundary rotation rate at the equator. The meridional circulation remains confined near the top of the radiative zone and deforms the poloidal magnetic field, which is almost confined within the radiative zone. When the inner core is assumed to be an electric insulator the equatorial region of the radiative zone rotates at 93% of the boundary rotation rate of the equator (figure 3, private communication).

Following the work of Garaud & Garaud we will investigate the following questions:

- What is the effect of the convective zone dynamics on the radiative zone? How does the rotational shear propagate into the radiative zone?
- Can we build a simple analytical magnetohydrodynamic model of the radiative zone to predict the rotation rate of the interior?
- How does the rotation rate vary when we impose different magnetic boundary conditions?

2 Governing equations

A sketch of our toy model is presented in figure 4. The region we are interested in mimics the solar radiative zone: a conducting fluid fills the shell between a uniformly rotating inner sphere and a differentially rotating outer sphere. An axial dipolar magnetic field is imposed; we present here the case of an open magnetic field. We make the following assumptions:

- The system has reached a steady state.
- The fluid is incompressible, and density is constant in the whole volume.

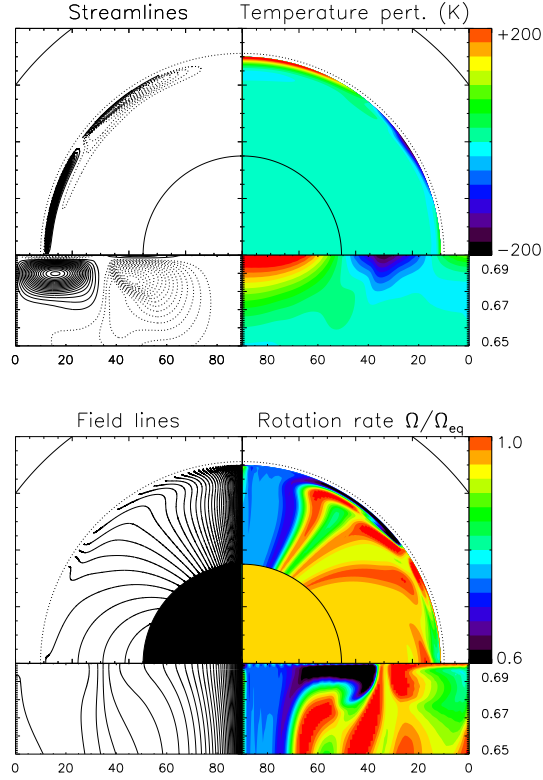


Figure 3: Results of a numerical simulation carried out with the code of Garaud & Garaud assuming an insulating inner core and a conducting outer boundary.

- In the solar radiative zone, the meridional circulation is slowed down by the strong stratification profile. Consequently, the angular momentum transported by the meridional flow is weak and will be neglected in this model. The non-linear terms are neglected.
- The imposed poloidal magnetic field is $\mathbf{B}_p = \nabla \times \left(\frac{A}{r \sin \theta} \hat{e}_\phi \right)$ with A the poloidal field potential $A(r, \theta) = B_0 r_{in}^3 \frac{\sin^2 \theta}{r}$ where B_0 is the imposed radial field at the poles on the inner sphere.

The magnetic field \mathbf{B} and the velocity field \mathbf{u} can therefore be decomposed in spherical coordinates as

$$\mathbf{B} = \left(\frac{1}{r^2 \sin \theta} \frac{\partial A}{\partial \theta}, -\frac{1}{r \sin \theta} \frac{\partial A}{\partial r}, B_\phi \right), \quad (3)$$

$$\mathbf{u} = (0, 0, r \sin \theta \Omega). \quad (4)$$

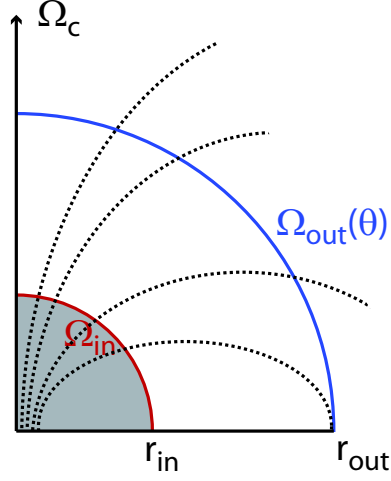


Figure 4: Diagram of our model.

The unknowns in our problem are the rotation rate Ω and the toroidal component of the magnetic field B_ϕ . The forms of Ω and B_ϕ are given by solving the azimuthal component of the Navier-Stokes equation and the magnetic induction equation respectively:

$$2\rho\Omega_c \times \mathbf{u} = -\nabla p + \left(\frac{1}{\mu_0}\nabla \times \mathbf{B}\right) \times \mathbf{B} + \rho\nu\nabla^2\mathbf{u}, \quad (5)$$

$$0 = \nabla \times (\mathbf{u} \times \mathbf{B}) - \eta\nabla \times \nabla \times \mathbf{B}, \quad (6)$$

where ρ is the density, Ω_c the (constant) global rotation of the system, μ_0 the magnetic permeability, ν and η the viscous and the magnetic diffusivities respectively.

Rigid rotation is assumed at the interface with the inner sphere. Therefore, the angular velocity at the inner boundary is

$$\Omega(r_{in}, \theta) = \Omega_{in}. \quad (7)$$

At the outer boundary we use the angular velocity profile of the solar convective zone

$$\Omega(r_{out}, \theta) = \Omega_{out}(\theta) = \Omega_{eq}(1 - a \cos^2 \theta - b \cos^4 \theta). \quad (8)$$

For the toroidal magnetic field we will test different boundary conditions in order to evaluate their influence on the rotation profile. If we assume that the region exterior to the fluid is an electric insulator then the toroidal magnetic field has to be zero at the boundaries:

$$B_\phi(r_{in}, \theta) = 0, \quad (9)$$

$$B_\phi(r_{out}, \theta) = 0. \quad (10)$$

If we assume that the fluid is surrounded by a conductor with the same permeability and conductivity as the fluid in the shell, then the radial and tangential components of the magnetic field and its derivatives must be continuous, and the induction equation in a steady state provides us the following boundary condition:

$$[\nabla^2 \mathbf{B}]_\phi = 0 \quad \text{for } r < r_{in} \quad \text{and} \quad r > r_{out}. \quad (11)$$

A difficult question is to decide which boundary condition is more relevant for the solar interior. This work will shed some light on the problem.

Finally the angular momentum must be conserved since the system is in a steady-state. Therefore the total torque applied in a spherical shell in the fluid is equal to zero:

$$\int_0^{\pi/2} \left(\rho \nu r^2 \sin^2 \theta \frac{\partial \Omega}{\partial r} + r \sin \theta \frac{B_r B_\phi}{\mu_0} \right) \sin \theta d\theta = 0. \quad (12)$$

The first term is the viscous torque whereas the second term is the magnetic torque. This equation determines uniquely Ω_{in} .

3 Insulating boundary conditions

We consider first the magnetic boundary conditions (9) and (10).

3.1 Solution in the bulk

In the bulk of the fluid, we will assume that we are in the limit of small magnetic and viscous diffusivities. This assumption is roughly valid in the Sun since the magnetic and viscous Ekman numbers E_η^\odot and E_ν^\odot which measure the ratio of magnetic and viscous diffusivities (resp.) over the Coriolis force are assumed to be about 3×10^{-14} and 2×10^{-15} (resp.). We will therefore neglect both the viscous and magnetic diffusivities. Given our assumptions, the equilibria that govern the system are

$$2\rho\Omega_c \times \mathbf{u} = -\nabla p + \left(\frac{1}{\mu_0} \nabla \times \mathbf{B} \right) \times \mathbf{B}, \quad (13)$$

$$0 = \nabla \times (\mathbf{u} \times \mathbf{B}). \quad (14)$$

Combining the ϕ -component of equation (14) and equation (4) and the fact that \mathbf{u} and \mathbf{B} are divergence-free, we obtain

$$[\nabla \times (\mathbf{u} \times \mathbf{B})]_\phi = r \sin \theta \mathbf{B}_p \cdot \nabla \Omega = 0. \quad (15)$$

Therefore in the bulk, the fluid is in a Ferraro iso-rotation state:

$$\mathbf{B}_p \cdot \nabla \Omega = 0, \quad (16)$$

or in other words Ω is constant along the poloidal magnetic field lines in a meridional plane. Using the ϕ -component of equation (13), we find

$$\left[\left(\frac{1}{\mu_0} \nabla \times \mathbf{B} \right) \times \mathbf{B} \right]_\phi = \frac{1}{r \sin \theta} \mathbf{B}_p \cdot \nabla (r \sin \theta B_\phi) = 0, \quad (17)$$

so the quantity $S \equiv r \sin \theta B_\phi$ is also constant along poloidal magnetic field lines. Moreover, we can show that $\mathbf{B}_p \cdot \nabla A = 0$. Since $A(r, \theta) \propto \frac{\sin^2 \theta}{r}$ then $\frac{\sin^2 \theta}{r}$ remains constant along a magnetic field line. This determines the relationship between θ_{in} (the colatitude of emergence of a field line out of the core) and θ_{out} the colatitude where the same field line connects with the convection zone (figure 5):

$$\frac{\sin^2 \theta_{in}}{r_{in}} = \frac{\sin^2 \theta_{out}}{r_{out}}. \quad (18)$$

The relationship between θ_{out} and θ_{in} can be written as $\theta_{out} = \Theta(\theta_{in})$ where

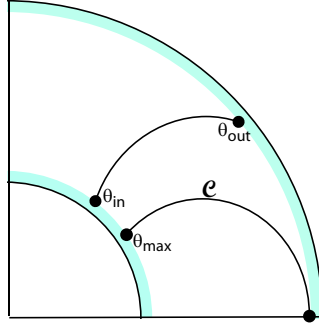


Figure 5: Picture of the dipolar magnetic field lines in the bulk of the fluid. The blue layers represent the boundary layers.

$$\Theta(\theta) = \sin^{-1} \left((r_{out}/r_{in})^{1/2} \sin \theta \right). \quad (19)$$

The magnetic field line which connects to the equatorial plane at r_{out} , which we call \mathcal{C} , crosses the inner sphere at the colatitude

$$\theta_{max} = \sin^{-1} \left(\frac{r_{in}}{r_{out}} \right)^{1/2} = \Theta \left(\frac{\pi}{2} \right). \quad (20)$$

The poloidal magnetic field lines which have a colatitude at the inner sphere between θ_{max} and $\pi/2$ connect to the inner sphere in the southern hemisphere without crossing the outer sphere. Because of isorotation we will assume that the equatorial region included between the equatorial plane and the line \mathcal{C} is rotating uniformly. Consequently the integration of the total torque in r_{in} can be reduced to the region above the magnetic field line \mathcal{C} :

$$\int_0^{\pi/2} \longrightarrow \int_0^{\theta_{max}}. \quad (21)$$

This approximation allows us to avoid the equatorial region where the boundary layer approach presented in the next section is not valid.

3.2 Solution in the boundary layers

Equations (5) and (6) in the azimuthal direction can be written in spherical coordinates as (Rüdiger & Kitchatinov, [3])

$$\eta \left[\frac{\partial}{\partial \theta} \left(\frac{1}{\sin \theta} \frac{\partial (B_\phi \sin \theta)}{\partial \theta} \right) + r \frac{\partial^2 (B_\phi r)}{\partial r^2} \right] = r \left(\frac{\partial \Omega}{\partial \theta} \frac{\partial A}{\partial r} - \frac{\partial \Omega}{\partial r} \frac{\partial A}{\partial \theta} \right), \quad (22)$$

$$\begin{aligned} \rho \nu \left[\frac{1}{\sin^3 \theta} \frac{\partial}{\partial \theta} \left(\sin^3 \theta \frac{\partial \Omega}{\partial \theta} \right) + \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^4 \frac{\partial \Omega}{\partial r} \right) \right] = \\ \frac{1}{\mu_0 r^2 \sin^3 \theta} \left(r \frac{\partial A}{\partial r} \frac{\partial (B_\phi \sin \theta)}{\partial \theta} - \sin \theta \frac{\partial A}{\partial \theta} \frac{\partial (B_\phi r)}{\partial r} \right). \end{aligned} \quad (23)$$

First we will consider the inner boundary layer. In order to simplify these equations we will use a stretched variable in the boundary layer:

$$\zeta = \frac{r - r_{in}}{\delta_{in}}. \quad (24)$$

We assume that variations with θ are very small compared to those in the ζ -direction. We will therefore neglect the θ -derivative in the equations. Note that this hypothesis happens to fail near the equator where the boundary layer thickness diverges (see below). Equations (22) and (23) become

$$\eta \frac{r_{in}}{\delta_{in}} \frac{\partial^2 B_\phi}{\partial \zeta^2} = - \frac{\partial \Omega}{\partial \zeta} \frac{\partial A}{\partial \theta}, \quad (25)$$

$$\rho \nu \mu_0 \frac{r_{in}^3 \sin^2 \theta}{\delta_{in}} \frac{\partial^2 \Omega}{\partial \zeta^2} = - \frac{\partial A}{\partial \theta} \frac{\partial B_\phi}{\partial \zeta}. \quad (26)$$

Combining these equations, we find

$$\frac{\partial^3 B_\phi}{\partial \zeta^3} = \frac{\partial B_\phi}{\partial \zeta}, \quad (27)$$

$$\frac{\partial^3 \Omega}{\partial \zeta^3} = \frac{\partial \Omega}{\partial \zeta}, \quad (28)$$

provided

$$\boxed{\delta_{in} = \frac{(\mu_0 \rho \nu \eta)^{1/2}}{2 \cos \theta B_0}}. \quad (29)$$

This boundary layer is an Hartmann layer. Its thickness is proportional to the geometric mean of the magnetic and viscous diffusivities. Moreover the layer becomes thinner if we increase the strength of the imposed magnetic field. The layer is singular near the equator where $\cos \theta \rightarrow 0$.

The solutions of (27) and (28) are

$$B_\phi(\zeta, \theta) = b_0^{in}(\theta) + b_1^{in}(\theta) e^\zeta + b_2^{in}(\theta) e^{-\zeta}, \quad (30)$$

$$\Omega(\zeta, \theta) = \Omega_0^{in}(\theta) + \Omega_1^{in}(\theta) e^\zeta + \Omega_2^{in}(\theta) e^{-\zeta}. \quad (31)$$

When $\zeta \rightarrow \infty$, Ω and B_ϕ have to remain bounded, and so $\Omega_1^{in} = b_1^{in} = 0$. Applying the boundary conditions and the matching conditions, we find

$$\Omega(\zeta \rightarrow 0, \theta) = \Omega_{in} = \Omega_0^{in}(\theta) + \Omega_2^{in}(\theta), \quad (32)$$

$$\Omega(\zeta \rightarrow \infty, \theta) = \Omega_0^{in}(\theta), \quad (33)$$

$$B_\phi(\zeta \rightarrow 0, \theta) = 0 = b_0^{in}(\theta) + b_2^{in}(\theta), \quad (34)$$

$$B_\phi(\zeta \rightarrow \infty, \theta) = b_0^{in}(\theta). \quad (35)$$

$\Omega_0^{in}(\theta)$ and $b_0^{in}(\theta)$ are the solutions in the bulk of the fluid where we neglect viscous and magnetic diffusivities. We will then call them $\Omega_0(r, \theta)$ and $b_0(r, \theta)$. Finally the solutions are

$$B_\phi(\zeta, \theta) = b_0(r_{in}, \theta)(1 - e^{-\zeta}), \quad (36)$$

$$\Omega(\zeta, \theta) = \Omega_0(r_{in}, \theta)(1 - e^{-\zeta}) + \Omega_{in}e^{-\zeta}. \quad (37)$$

Using either (25) or (26), we find the relation between b_0 and Ω_0 :

$$b_0(\theta, r_{in}) = \frac{\delta_{in}}{\eta r_{in}} \frac{\partial A}{\partial \theta} (\Omega_0(r_{in}, \theta) - \Omega_{in}) = \left(\frac{\mu_0 \rho \nu}{\eta} \right)^{1/2} r_{in} \sin \theta (\Omega_0(r_{in}, \theta) - \Omega_{in}). \quad (38)$$

Since the boundary condition on the inner core imposed that $B_\phi(r_{in}, \theta) = 0$, and $\rho \nu$ is assumed to be constant in the boundary layer, (12) becomes

$$\int_0^{\theta_{max}} \sin^3 \theta \frac{\partial \Omega}{\partial r} d\theta = 0, \quad (39)$$

which can be written as

$$\int_0^{\theta_{max}} \frac{\sin^3 \theta}{\delta_{in}} \frac{\partial \Omega}{\partial \zeta} d\theta = 0. \quad (40)$$

Using (37), we have

$$\frac{\partial \Omega}{\partial \zeta} = (\Omega_0(r_{in}, \theta) - \Omega_{in}) e^{-\zeta} \implies \left. \frac{\partial \Omega}{\partial \zeta} \right|_{\zeta=0} = \Omega_0(r_{in}, \theta) - \Omega_{in}. \quad (41)$$

Moreover

$$\delta_{in} \propto \frac{1}{\cos \theta}. \quad (42)$$

Therefore (40) becomes

$$\int_0^{\theta_{max}} \sin^3 \theta \cos \theta (\Omega_0(r_{in}, \theta) - \Omega_{in}) d\theta = 0. \quad (43)$$

Condition (12) should be valid everywhere in the fluid. In the bulk of the fluid, the kinematic viscosity is neglected, so the torque applied to a spherical shell at radius $r_{in} + \epsilon$, where $\epsilon > \delta_{in}$, is only the magnetic torque and (12) becomes

$$\int_0^{\pi/2} (r_{in} + \epsilon) \sin^2 \theta \frac{B_r B_\phi(r_{in} + \epsilon, \theta)}{\mu_0} d\theta = 0. \quad (44)$$

Using the value of B_ϕ in the bulk given by (38) and since $B_r \propto \cos \theta$ we find the condition

$$\int_0^{\pi/2} \sin^3 \theta \cos \theta (\Omega_0(r_{in} + \epsilon, \theta) - \Omega_{in}) d\theta = 0, \quad (45)$$

which is the same condition as previously. The viscous torque acts only on the boundary layer whereas the magnetic torque acts in the bulk, but together they maintain the previous condition everywhere in the fluid.

We have to determine the value of $\Omega_0(r_{in}, \theta)$ in order to calculate a value for Ω_{in} . To match $\Omega_0(r_{in}, \theta)$ to its value at the outer boundary we have to find the solution in the outer boundary.

We use the stretched variable in the outer boundary layer:

$$\xi = \frac{r_{out} - r}{\delta_{out}}. \quad (46)$$

Proceeding as for the inner boundary layer, we find

$$B_\phi(\xi, \theta) = b_0(r_{out}, \theta)(1 - e^{-\xi}), \quad (47)$$

$$\Omega(\xi, \theta) = \Omega_0(r_{out}, \theta)(1 - e^{-\xi}) + \Omega_{out}(\theta)e^{-\xi}, \quad (48)$$

with

$$b_0(r_{out}, \theta) = \left(\frac{\rho \nu \mu_0}{\eta} \right)^{1/2} r_{out} \sin \theta (\Omega_{out}(\theta) - \Omega_0(r_{out}, \theta)), \quad (49)$$

and

$$\delta_{out} = \frac{(\mu_0 \rho \nu \eta)^{1/2}}{2 \cos \theta B_0} \left(\frac{r_{out}}{r_{in}} \right)^3 = \delta_{in} \left(\frac{r_{out}}{r_{in}} \right)^3. \quad (50)$$

The thickness of the outer boundary layer δ_{out} is larger than that of the inner boundary because the imposed magnetic field is proportional to $1/r^3$, and therefore weaker at the outer boundary.

Since $S = B_\phi r \sin \theta$ is constant along a poloidal magnetic field line, we have

$$b_0(r_{in}, \theta_{in}) r_{in} \sin \theta_{in} = b_0(r_{out}, \theta_{out}) r_{out} \sin \theta_{out}. \quad (51)$$

Using (38) and (49), we find

$$(r_{in} \sin \theta)^2 (\Omega_0(r_{in}, \theta) - \Omega_{in}) = (r_{out} \sin \Theta(\theta))^2 (\Omega_{out}(\Theta(\theta)) - \Omega_0(r_{out}, \Theta(\theta))) . \quad (52)$$

Since the bulk of the fluid is in a state of iso-rotation, we expect that $\Omega_0(r_{in}, \theta) = \Omega_0(r_{out}, \Theta(\theta))$. Consequently,

$$\Omega_0(r_{in}, \theta) = \frac{1}{1 + \frac{(r_{in} \sin \theta)^2}{(r_{out} \sin \Theta(\theta))^2}} \Omega_{out}(\Theta(\theta)) + \frac{1}{1 + \frac{(r_{out} \sin \Theta(\theta))^2}{(r_{in} \sin \theta)^2}} \Omega_{in}, \quad (53)$$

$$\Omega_0(r_{in}, \theta) = \frac{1}{1 + \frac{r_{in}^3}{r_{out}^3}} \Omega_{out}(\Theta(\theta)) + \frac{1}{1 + \frac{r_{out}^3}{r_{in}^3}} \Omega_{in}. \quad (54)$$

The condition (43) becomes

$$\int_0^{\theta_{max}} \sin^3 \theta \cos \theta \left[\frac{1}{1 + \frac{r_{in}^3}{r_{out}^3}} \Omega_{out}(\theta_{out}) + \Omega_{in} \left(-1 + \frac{1}{1 + \frac{r_{out}^3}{r_{in}^3}} \right) \right] d\theta = 0, \quad (55)$$

with

$$\Omega_{out}(\Theta(\theta)) = \Omega_{eq} \left[1 - a \left(1 - \frac{r_{in}}{r_{out}} \sin^2 \theta \right) - b \left(1 - \frac{r_{in}}{r_{out}} \sin^2 \theta \right)^2 \right]. \quad (56)$$

We find

$$\frac{\Omega_{in}}{\Omega_{eq}} = \frac{\int_0^{\theta_{max}} \left(1 - a \left(1 - \frac{r_{in}}{r_{out}} \sin^2 \theta \right) - b \left(1 - \frac{r_{in}}{r_{out}} \sin^2 \theta \right)^2 \right) \sin^3 \theta \cos \theta d\theta}{\int_0^{\theta_{max}} \sin^3 \theta \cos \theta d\theta}, \quad (57)$$

$$\frac{\Omega_{in}}{\Omega_{eq}} = 1 - \frac{a}{3} - \frac{b}{6}. \quad (58)$$

Using the values provided by the observations of the Sun, $a = 0.17$ and $b = 0.08$, we find $\Omega_{in}/\Omega_{eq} = 0.93$. The remarks that we can draw from this result are:

- The result does not depend on the gap between the two spheres.
- It is the same value as that observed in the Sun. Therefore we can wonder if this value has a physical explanation. It is hard to answer this question but it is possible that our simple model catches an essential feature of the dynamics of the solar radiative zone.

3.3 No inner boundary

We wonder what happens if there is no inner boundary layer. Indeed the interface between the nuclear core and the radiative zone is not a physical barrier submitted to no-slip conditions. The nuclear core is defined simply as the place where the temperature is large enough to trigger nuclear reactions. Therefore the interface between the inner core and the radiative zone is merely an isotherm, and not a physical boundary.

The total torque applied on the outer sphere is

$$\int_0^{\pi/2} \rho \nu r_{out}^2 \sin^3 \theta \left. \frac{\partial \Omega}{\partial r} \right|_{r_{out}} d\theta = 0, \quad (59)$$

$$\left. \frac{\partial \Omega}{\partial r} \right|_{r_{out}} = -\frac{1}{\delta_{out}} \left. \frac{\partial \Omega}{\partial \xi} \right|_{\xi \rightarrow 0} = -\frac{1}{\delta_{out}} (\Omega_0(r_{out}, \theta) - \Omega_{out}(\theta)). \quad (60)$$

We will match the solution in the bulk at r_{out} with Ω_{in} the angular velocity at r_{in} . Under these assumption the whole radiative zone is rotating uniformly at the angular velocity Ω_{in} (see figure 6). Consequently we have

$$\int_0^{\pi/2} \sin^3 \theta \cos \theta (\Omega_{out}(\theta) - \Omega_{in}) d\theta = 0, \quad (61)$$

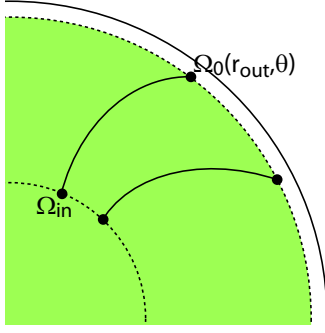


Figure 6: Diagram of the model when considering no inner boundary. The radiative zone is uniformly rotating and the outer boundary layer undergoes a strong shear.

where

$$\Omega_{out}(\theta) = \Omega_{eq} (1 - a \cos^2 \theta - b \cos^4 \theta). \quad (62)$$

The whole interior is rotating at the angular velocity

$$\frac{\Omega_{in}}{\Omega_{eq}} = \frac{\int_0^{\pi/2} \sin^3 \theta \cos \theta (1 - a \cos^2 \theta - b \cos^4 \theta) d\theta}{\int_0^{\pi/2} \sin^3 \theta \cos \theta d\theta}, \quad (63)$$

$$\frac{\Omega_{in}}{\Omega_{eq}} = 1 - \frac{a}{3} - \frac{b}{6}. \quad (64)$$

We find the same expression as previously. If a mechanism is able to impose an uniform rotation to the radiative zone interior, then the presence of the outer boundary alone prevents the transfer of the angular velocity from the convective zone to the deep radiative zone along open magnetic field lines. In the outer boundary layer the zonal shear is very strong, as in the solar tachocline.

3.4 Characteristic features of this toy model of the solar radiative zone

Our boundary layer analysis provided a solution valid from the poles to the magnetic field line \mathcal{C} . The solution in the equatorial zone can be deduced from geometrical considerations:

- The solutions are either symmetric or antisymmetric with respect to the equatorial plane due to the symmetry of the imposed poloidal magnetic field. The angular velocity, for example, is symmetric.
- The toroidal field on the other hand is antisymmetric with respect to the equatorial plane; therefore S is also antisymmetric. An important consequence of this fact is that since S has to be constant along the magnetic field lines, S and so the toroidal field must be equal to zero in the equatorial region where field lines connect the northern and southern hemispheres.

- Another consequence is that the equatorial region, defined as the region between the equatorial plane and the magnetic field line \mathcal{C} , rotates at the constant angular velocity $\Omega_{in} = 0.93\Omega_{eq}$.
- The presence of a rapidly rotating layer around the equatorial region is due to the fact that a magnetic field line that connects just above the equator in the outer sphere transfers the fast angular velocity of the convective zone at the equator until the inner sphere.
- This rapidly rotating layer shears the poloidal magnetic field lines and creates a strong toroidal magnetic field (the Ω -effect).
- Towards the poles the angular velocity transferred from the convective zone is lower. Therefore the angular velocity in the bulk decreases at high latitudes.

An outline of the geometry of our asymptotic model of the solar radiative zone is presented figure 7.

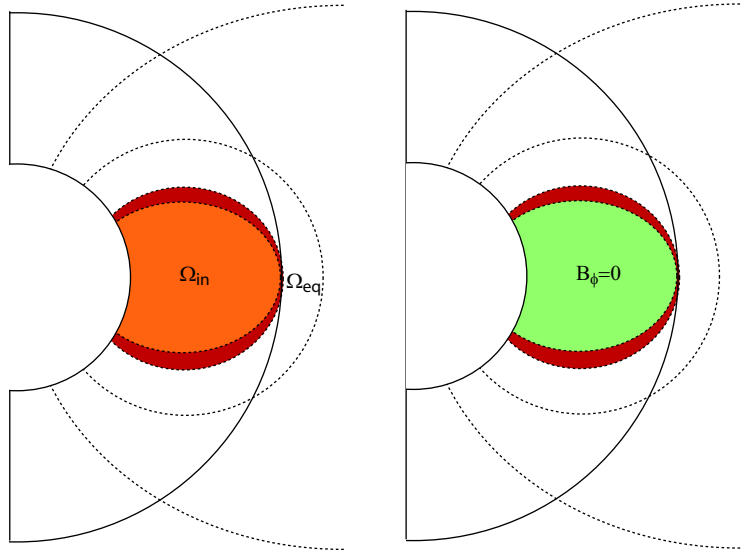


Figure 7: Asymptotic model of the conducting fluid flow between two rotating spheres (no viscosity, no magnetic diffusivity)

3.5 Comparison with numerical results

The code used to compare the analytical results with numerical simulations is a stripped version of the code of Garaud & Garaud [1] modified to solve equations (22) and (23) only. The same assumptions as in our analytical model are made except that the kinematic and

magnetic diffusivities are not neglected in the main body of the fluid. The same boundary conditions are adopted. Numerical simulations are limited by the fact that assuming very low diffusivities means that the boundary layers are very thin and can not be resolved numerically. To overcome this problem, diffusivities are enhanced by multiplying them by a factor f . $f = 1$ means that the values adopted in the simulations are the solar values. Figure 8 show the convergence of the numerical value of the ratio Ω_{in}/Ω_{eq} towards the analytical value of 0.93 when decreasing the factor f . The results of a numerical simulation are presented in figure 9. The structure of the angular velocity and the toroidal magnetic field are similar to the ones predicted by our model in section 3.4. The rapidly rotating layer is visible.

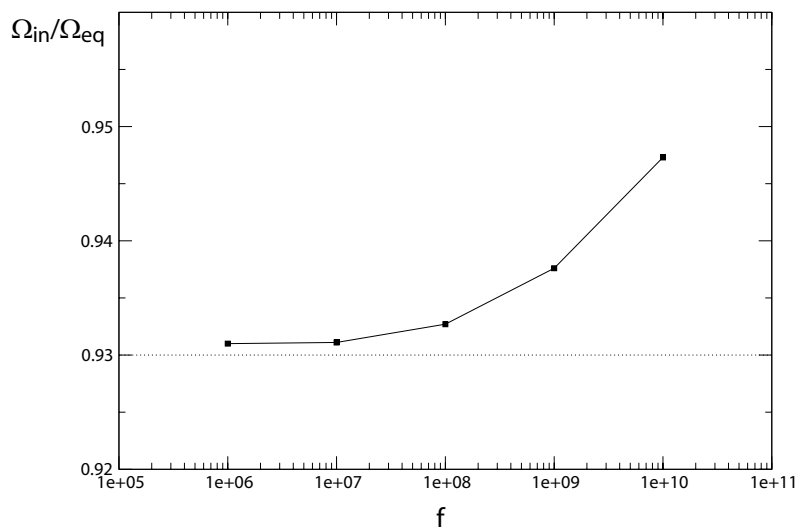


Figure 8: Ratio of the angular velocity at the inner sphere and at the equator of the convective zone in function of the factor f . The dots represent the results of the numerical simulations and the dashed line the analytical value. Note that $f = 10^{10}$ corresponds to $E_\nu = 2 \cdot 10^{-5}$ and $E_\eta = 3 \cdot 10^{-4}$.

The total torque (12) computed numerically as a function of the colatitude (figure 10) is equal to zero if $\theta > \theta_{max}$. This remark validates our assumption of neglecting the contribution of the equatorial region (section 3.1).

4 Conducting boundary conditions

We now consider applying boundary condition (11) to the toroidal magnetic field. In the convective zone and inner core the media are assumed to have the same electric conductivity and permeability as the fluid in the shell. Therefore the tangential components of the magnetic field are continuous. Outside the fluid domain B_ϕ satisfies the induction equation

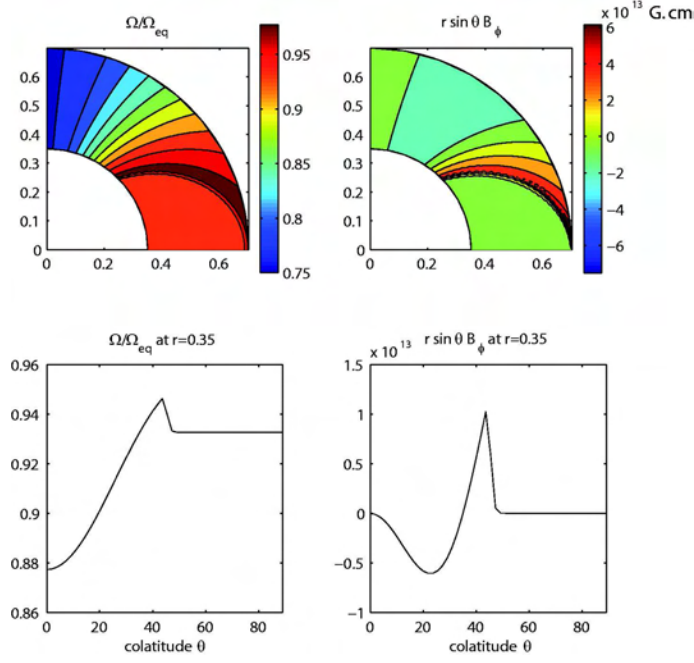


Figure 9: Results of the numerical simulations computed at $f = 10^8$. Left: angular velocity in a quadrant of the meridional plane (top) and at a fixed radius just above the inner boundary layer (bottom). Right: function $S = r \sin \theta B_\phi$ similarly plotted.

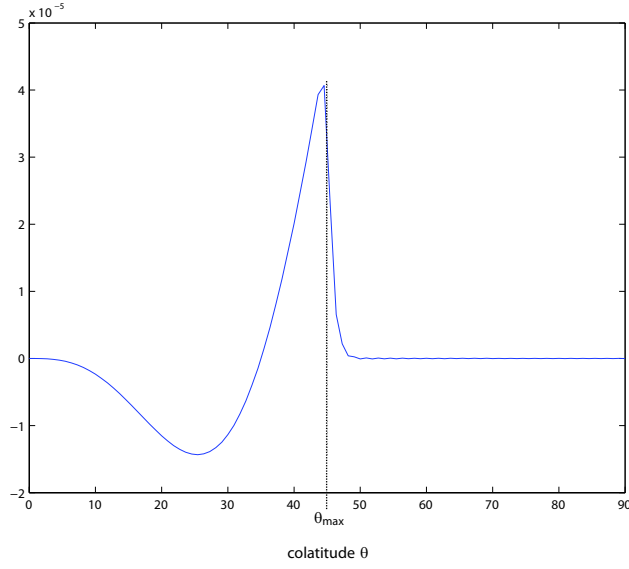


Figure 10: Total torque at r_{in} in function of the colatitude.

in the steady-state:

$$[\nabla^2 \mathbf{B}]_\phi = 0 \quad (65)$$

$$\Rightarrow \quad \nabla^2 B_\phi - \frac{B_\phi}{r^2 \sin^2 \theta} = 0. \quad (66)$$

The solution of this equation on both sides of the cavity are

$$B_\phi(r \leq r_{in}, \theta) = \sum_{n=0}^{\infty} a_n^{in} r^n \frac{d}{d\theta} [P_n(\cos \theta)], \quad (67)$$

$$B_\phi(r \geq r_{out}, \theta) = \sum_{n=0}^{\infty} a_n^{out} r^{-(n+1)} \frac{d}{d\theta} [P_n(\cos \theta)], \quad (68)$$

where P_n is the n -th Legendre polynomial. The continuity of B_ϕ and its r -derivative has to be verified across the interface. Applying these conditions to the solutions in the boundary layers (30), we find

$$\sum_{n=0}^{\infty} a_n^{in} r_{in}^n \frac{d}{d\theta} [P_n(\cos \theta)] = b_0(r_{in}, \theta) + b_2^{in}(\theta), \quad (69)$$

$$\sum_{n=0}^{\infty} a_n^{in} n r_{in}^{n-1} \frac{d}{d\theta} [P_n(\cos \theta)] = -\frac{1}{\delta_{in}} b_2^{in}(\theta), \quad (70)$$

$$\sum_{n=0}^{\infty} a_n^{out} r_{out}^{-(n+1)} \frac{d}{d\theta} [P_n(\cos \theta)] = b_0(r_{out}, \theta) + b_2^{out}(\theta), \quad (71)$$

$$\sum_{n=0}^{\infty} -a_n^{out} (n+1) r_{out}^{-(n+2)} \frac{d}{d\theta} [P_n(\cos \theta)] = \frac{1}{\delta_{out}} b_2^{out}(\theta). \quad (72)$$

The system of equations arising from this choice of boundary conditions is too complex to solve in spherical coordinates. For simplicity, we adopt a local cartesian box model for the boundary layers.

4.1 The Cartesian box

The model is illustrated in figure 11. We make the same assumptions as previously and introduce periodicity in the y -direction (equivalent to the θ -direction in spherical coordinates). The system is invariant in the x -direction, which corresponds to the ϕ -direction in the original spherical coordinates. A magnetic field is imposed in the z -direction, and depends only on y . The magnetic field and the velocity are decomposed in cartesian coordinates as

$$\mathbf{B} = (b, 0, B_0), \quad (73)$$

$$\mathbf{u} = (u, 0, 0). \quad (74)$$

We solve the system for the unknown b and u . The Navier-Stokes and the induction equations in the x -direction become under our assumptions

$$B_0(y) \frac{\partial b}{\partial z} = -\rho \nu \mu_0 \nabla^2 u, \quad (75)$$

$$B_0(y) \frac{\partial u}{\partial z} = -\eta \nabla^2 b. \quad (76)$$

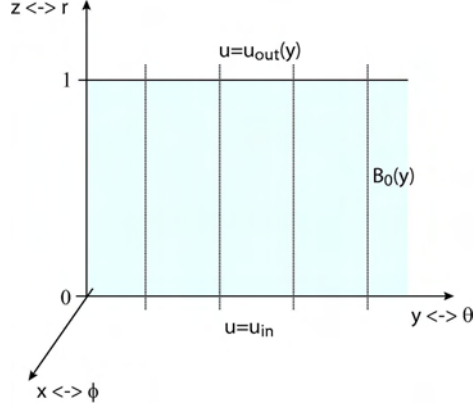


Figure 11: Diagram of the cartesian box.

4.2 Boundary layers

As in the spherical case, we use a stretched variable in the boundary layers. For the inner boundary, we use

$$\zeta = \frac{z}{\delta_{in}}, \quad (77)$$

and for the outer boundary

$$\xi = \frac{1-z}{\delta_{out}}. \quad (78)$$

In our boundary layer approximation, the y -derivative will be neglected. The previous equations become in the inner boundary

$$B_0(y) \frac{\partial b}{\partial \zeta} = -\frac{\rho \nu \mu_0}{\delta_{in}} \frac{\partial^2 u}{\partial \zeta^2}, \quad (79)$$

$$B_0(y) \frac{\partial u}{\partial \zeta} = -\frac{\eta}{\delta_{in}} \frac{\partial^2 u}{\partial \zeta^2}. \quad (80)$$

In the outer boundary layer the equations are similar apart from the sign. Combining the above equations in each boundary layer, we find

$$\frac{\partial^3 b}{\partial \zeta^3} = \frac{\partial b}{\partial \zeta}, \quad (81)$$

$$\frac{\partial^3 u}{\partial \zeta^3} = \frac{\partial u}{\partial \zeta}, \quad (82)$$

with

$$\delta_{in} = \delta_{out} = \frac{(\rho \nu \eta \mu_0)^{1/2}}{B_0(y)} = \delta. \quad (83)$$

As previously, these boundary layers are Hartmann layers. The solutions of these equations are, in the inner boundary,

$$u(\zeta, y) = u_0(z = 0, y) + u_2^{in}(y)e^{-\zeta}, \quad (84)$$

$$b(\zeta, y) = b_0(z = 0, y) + b_2^{in}(y)e^{-\zeta}, \quad (85)$$

$$b_2^{in}(y) = \frac{\rho\nu\eta}{\delta B_0(y)} u_2^{in}(y). \quad (86)$$

The solutions in the outer boundary are analogous, apart from sign changes. Angular momentum conservation gives us the condition

$$\int_0^{2\pi} \left[\rho\nu \frac{\partial u}{\partial z} + \frac{1}{\mu_0} B_0 b \right] dy = 0. \quad (87)$$

Moreover, continuity of the velocity in the bulk along the magnetic field lines² of B_0 provides another condition

$$u_0(z = 1, y) = u_0(z = 0, y), \quad (88)$$

$$b_0(z = 1, y) = b_0(z = 0, y). \quad (89)$$

The boundary conditions on the velocity are

$$u(z = 0, y) = u_{in}, \quad (90)$$

$$u(z = 1, y) = u_{out}(y). \quad (91)$$

In order to find an analogy with the previous case, we will first study the case where $b = 0$ outside on both sides.

4.3 Insulating boundaries

In this case, the boundary conditions for the tangential magnetic field are

$$b(z = 0, y) = b_0(z = 0, y) + b_2^{in}(y) = 0, \quad (92)$$

$$b(z = 1, y) = b_0(z = 1, y) + b_2^{out}(y) = 0. \quad (93)$$

Using (86), we find

$$2b_0(z, y) = -b_2^{in} - b_2^{out} = \frac{\rho\nu\eta}{\delta B_0(y)} (u_2^{out}(y) - u_2^{in}(y)). \quad (94)$$

When $\zeta \rightarrow 0$ and $\xi \rightarrow 0$,

$$u_0(z = 0, y) + u_2^{in}(y) = u_{in}, \quad (95)$$

$$u_0(z = 0, y) + u_2^{out}(y) = u_{out}(y), \quad (96)$$

and using the continuity of u_0 in the bulk along the y -lines

$$u_2^{in}(y) - u_2^{out}(y) = u_{in} - u_{out}(y). \quad (97)$$

Plugging this equation into (87) we find

$$\int_0^{2\pi} B_0(y) (u_{out}(y) - u_{in}) dy = 0. \quad (98)$$

This equation is analogous to (43) by geometrical transformation.

²lines of constant y .

4.4 Conducting boudaries

When conducting boundaries are applied, the tangential component of the field is continuous across the interface. Outside of the fluid, the tangential magnetic field satisfies

$$\frac{\partial^2 b}{\partial y^2} + \frac{\partial^2 b}{\partial z^2} = 0. \quad (99)$$

For $z \notin [0, 1]$ the field can be written as

$$b(z \leq 0, y) = \beta_0^{in} + \sum_{n=1}^{\infty} \beta_n^{in}(z) e^{iny}, \quad (100)$$

$$b(z \geq 1, y) = \beta_0^{out} + \sum_{n=1}^{\infty} \beta_n^{out}(z) e^{iny}, \quad (101)$$

where the functions β_n satisfies

$$-n^2 \beta_n + \frac{\partial^2 \beta_n}{\partial z^2} = 0. \quad (102)$$

The solution of this equation is

$$\beta_n(z) = \alpha_n e^{\pm nz}, \quad (103)$$

which has to remain bounded

$$\text{if } z > 1 : \quad \beta_n^{out}(z) = \alpha_n^{out} e^{-nz}, \quad (104)$$

$$\text{if } z < 0 : \quad \beta_n^{in}(z) = \alpha_n^{in} e^{nz}. \quad (105)$$

The field is continuous at the interface so

$$b_0^{in}(y) + b_2^{in}(y) = \beta_0^{in} + \sum \beta_n^{in}(z=0) e^{iny}, \quad (106)$$

$$b_0^{out}(y) + b_2^{out}(y) = \beta_0^{out} + \sum \beta_n^{out}(z=1) e^{iny}. \quad (107)$$

The z -derivative of b also has to be continuous across the interface which yields the two additional equations

$$-\frac{1}{\delta} b_2^{in}(y) = \sum_{n=1}^{\infty} n \beta_n^{in}(z=0) e^{iny}, \quad (108)$$

$$\frac{1}{\delta} b_2^{out}(y) = \sum_{n=1}^{\infty} -n \beta_n^{out}(z=1) e^{iny}. \quad (109)$$

Adding the previous equations yields

$$\frac{1}{\delta} (b_2^{in}(y) + b_2^{out}(y)) = - \sum_{n=1}^{\infty} n (\beta_n^{in}(z=0) + \beta_n^{out}(z=1)) e^{iny}. \quad (110)$$

Combining the relation between b_2 and u_2 in both boundary layers (86), we find

$$b_2^{in}(y) + b_2^{out}(y) = \frac{\rho\eta\nu}{B_0(y)} \frac{1}{\delta} (u_2^{in} - u_2^{out}), \quad (111)$$

$$b_2^{in}(y) + b_2^{out}(y) = \frac{\rho\eta\nu}{B_0(y)} \frac{1}{\delta} (u_{in} - u_{out}(y)). \quad (112)$$

The results of the previous equations gives us

$$\frac{\rho\eta\nu}{B_0(y)} \frac{1}{\delta^2} (u_{in} - u_{out}(y)) = - \sum_{n=1}^{\infty} n (\beta_n^{in}(z=0) + \beta_n^{out}(z=1)) e^{iny} \quad (113)$$

$$\begin{aligned} \Rightarrow \int_0^{2\pi} \frac{\rho\eta\nu}{B_0(y)} \frac{1}{\delta^2} (u_{in} - u_{out}(y)) dy = \\ - \int_0^{2\pi} \sum_{n=1}^{\infty} n (\beta_n^{in}(z=0) + \beta_n^{out}(z=1)) e^{iny} dy. \end{aligned} \quad (114)$$

Since the box is 2π periodic in the y -direction and $\beta_n^{in}(z=0)$ and $\beta_n^{out}(z=1)$ are constant, the right-hand side is equal to zero. Since $\delta \propto 1/B_0$ we have

$$\int_0^{2\pi} B_0(y) (u_{out}(y) - u_{in}) dy = 0. \quad (115)$$

This condition is similar to the case where $b = 0$ outside. Therefore we expect to find the same rotation rate in the radiative zone with this kind of boundary condition. This means that the boundary conditions imposed on the tangential field should *a priori* not influence the rotation rate of the fluid.

4.5 Numerical results

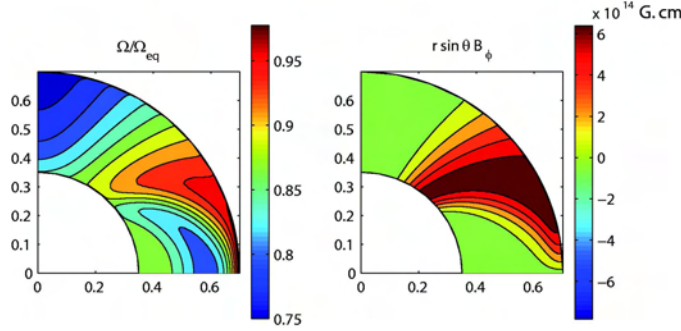


Figure 12: Numerical results with conducting boundary conditions.

The results of a simulation where $[\nabla^2 B]_\phi = 0$ is imposed on both sides is presented in figure 12. We find that the inner core rotates with angular velocity $\Omega_{in} = 0.86\Omega_{eq}$, contradicting the results of our analytic model of section 4.4. This value is lower than

the case with insulating boundaries because we notice that the fast angular velocity of the bottom of the equatorial convective zone is not propagated until the inner region of the radiative zone. We observe that the toroidal field B_ϕ is about ten times stronger than in the case where $B_\phi = 0$ was imposed on both boundaries. We also observe, in figure 12, that the Ferraro state is broken in the bulk of the fluid:

$$\mathbf{B}_p \cdot \nabla \Omega = \eta \nabla^2 B_\phi \neq 0. \quad (116)$$

It therefore appears that magnetic diffusivity plays a role in the bulk of the fluid which explains why our analytical model fails to reproduce the numerical results.

5 Conclusions

This work allows us to draw several conclusions and perspectives:

- Our toy model with an open magnetic field and vacuum magnetic boundary conditions provides a prediction of the interior rotation rate of the radiative zone in good agreement with both solar observations and full numerical simulations (including meridional circulations and all the non-linear terms (figure 3)). We find the solar value whereas the structure is not uniformly rotating. Moreover it allows us to understand the force balance of the system. The interior rotation rate of our simple model can provide insight into the physical processes that occur in the solar radiative zone.
- The boundary conditions chosen for the toroidal field have important consequences for the rotation of the fluid. However, the more relevant conditions for the Sun are difficult to determine.
- The confined field case needs to be investigated because of its believed relevance to the Sun.

References

- [1] P. GARAUD AND J.-D. D. GARAUD, *Dynamics of the solar tachocline II: the stratified case*, ArXiv e-prints, 806 (2008).
- [2] D. GOUGH AND M. E. MCINTYRE, *Inevitability of a magnetic field in the Sun's radiative interior*, Nature, 394 (1998), pp. 755–757.
- [3] G. RUDIGER AND L. L. KITCHATINOV, *The slender solar tachocline: a magnetic model*, Astronomische Nachrichten, 318 (1997), pp. 273–279.
- [4] J. SCHOU, H. M. ANTIA, S. BASU, R. S. BOGART, R. I. BUSH, S. M. CHITRE, J. CHRISTENSEN-DALSGAARD, M. P. DI MAURO, W. A. DZIEMBOWSKI, A. EFF-DARWICH, D. O. GOUGH, D. A. HABER, J. T. HOEKSEMA, R. HOWE, S. G. KORZENNIK, A. G. KOSOVICHEV, R. M. LARSEN, F. P. PIJPERS, P. H. SCHERRER, T. SEKII, T. D. TARBELL, A. M. TITLE, M. J. THOMPSON, AND J. TOOMRE, *Heliioseismic Studies of Differential Rotation in the Solar Envelope by the Solar Oscillations Investigation Using the Michelson Doppler Imager*, The Astrophysical Journal, 505 (1998), pp. 390–417.

Bounds on Surface Stress Driven Flows

George Hagstrom
Charlie Doering

January 12, 2009

1 Introduction

In the past fifteen years the background method of Constantin, Doering, and Hopf has been used to derive rigorous bounds on transport in fluid mechanics equations. The background method is based on the decomposition of the velocity or temperature field into a background field and a fluctuation field. The background field is specified so that the fluctuation field satisfies a partial differential equation with homogeneous boundary conditions. Many bounding problems have a variational formulation under the background method. The background method has been used to prove bounds on the mechanical dissipation rate in various driven turbulent flows. This work was started by Constantin and Doering [4]. They studied the mechanical dissipation in shear flows at high Reynolds number. Since then there has been a large number of papers about dissipation in turbulent shear flow with different boundary conditions and geometries.

We extend this technique to flows driven by stress at the boundary, namely Marangoni convection and surface shear stress driven flow. In the case of surface shear stress driven flow we study the energy stability of the laminar flow solution as a function of the Grashoff number and prove an upper bound on the friction coefficient for high Reynolds number. Tang, Caulfield, and Young [6] first used the background method to prove bounds for this type of problem and we compare our results to theirs.

The problem of Marangoni convection is related to that of stress driven shear flow. In the Marangoni case the surface stress is due to surface tension gradients. Pearson [5] developed the linear theory for Marangoni convection in 1958, and Davis [3] used variational methods to study the nonlinear stability problem in 1969. In the case of infinite Prandtl number we use nonvariational methods to improve the estimate of the critical Marangoni number for nonlinear stability of the conduction solution. We also use the background method to prove an upper bound on the Nusselt number in Marangoni convection.

2 Surface Stress Driven Flow

We consider the problem of flow in the two dimensional domain in Figure 2 subject to periodic boundary conditions in the horizontal x direction, no slip conditions at $z = 0$, and

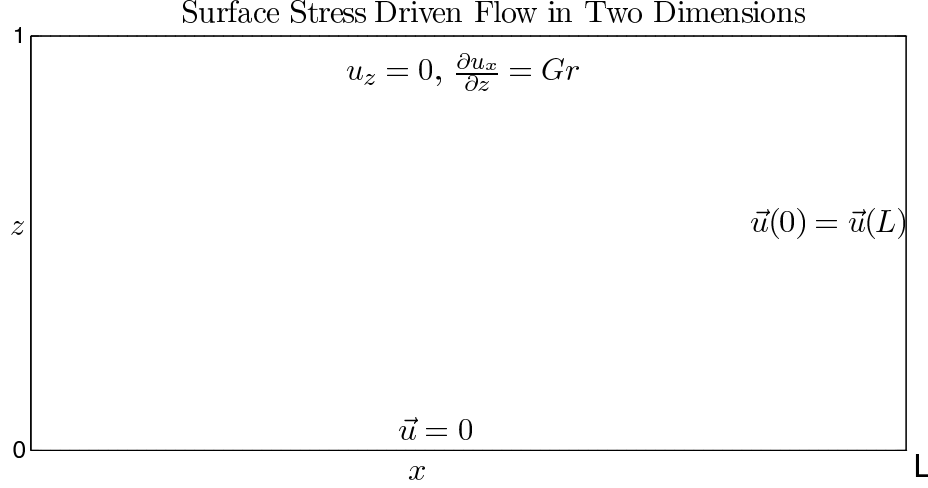


Figure 1: Domain for the surface stress driven flow.

a fixed shear stress at $z = h$. The full system of equations is:

$$\frac{\partial u}{\partial t} + u \cdot \nabla u + \nabla p = \nu \Delta u \quad (1)$$

$$\nabla \cdot u = 0 \quad (2)$$

$$u|_{z=0} = 0 \quad (3)$$

$$u_z|_{z=h} = 0 \quad (4)$$

$$\nu \frac{\partial u_x}{\partial z} = \tau \quad (5)$$

We take periodic boundary conditions in x , with a domain length L .

2.1 Energy Stability Analysis in Two Dimensions

The laminar flow solution of these equations is $u_x = \frac{\tau z}{\nu}$. We perform the energy stability analysis of this solution by making the substitution $u = U + \frac{\tau z}{\nu} \hat{x}$. The resulting system of equations is:

$$\frac{\partial U}{\partial t} + U \cdot \nabla U + \frac{\tau z}{\nu} \frac{\partial U}{\partial x} + U_z \frac{\tau}{\nu} \hat{x} + \nabla p = \nu \Delta U \quad (6)$$

$$\nabla \cdot U = 0 \quad (7)$$

$$U_x|_{z=0} = 0 \quad (8)$$

$$\frac{\partial U_x}{\partial z}|_{z=h} = 0 \quad (9)$$

$$U_z|_{z=0,1} = 0 \quad (10)$$

Take the dot product of the momentum equation with U and integrate over the domain to derive the energy expression:

$$\frac{1}{2} \frac{d\|U\|^2}{dt} + \int \frac{\tau}{\nu} U_x U_z dx dz = -\nu \|\nabla U\|^2 \quad (11)$$

If $\frac{d\|U\|^2}{dt} < 0$ for all U satisfying the perturbation's boundary conditions then the base solution is stable. This derivative is given by a quadratic form $-Q$ in U :

$$Q = \nu \|\nabla U\|^2 + \int \frac{\tau}{\nu} U_x U_z dx dz \quad (12)$$

We use variational methods to minimize Q , subject to the constraints $\nabla \cdot U = 0$ and $\|U\|^2 = 1$. The resulting Euler-Lagrange equations are:

$$\lambda U_x = -\nu \Delta U_x + \frac{\partial q}{\partial x} + \frac{\tau}{2\nu} U_z \quad (13)$$

$$\lambda U_z = -\nu \Delta U_z + \frac{\partial q}{\partial z} + \frac{\tau}{2\nu} U_x \quad (14)$$

$$\nabla \cdot U = 0 \quad (15)$$

$$\|U\|^2 = 1 \quad (16)$$

Here q is the Lagrange multiplier associated with incompressibility and λ is the Lagrange multiplier associated with the normalization condition. We non-dimensionalize the equations by making the substitutions $x' = \frac{x}{h}$ and $z' = \frac{z}{h}$. Define the Grashoff number $Gr = \frac{\tau h^2}{\nu^2}$. Then if $U'(x', z') = U(x, z)$ we have the following equations:

$$\lambda U_x = -\Delta U_x + \frac{Gr}{2} U_z + \frac{\partial q}{\partial x} \quad (17)$$

$$\lambda U_z = -\Delta U_z + \frac{Gr}{2} U_x + \frac{\partial q}{\partial z} \quad (18)$$

$$(19)$$

If the smallest eigenvalue $\lambda_{min} > 0$ then Q is positive definite and the base solution is stable. Since the system is two dimensional we introduce the stream function Ψ , which satisfies $\frac{\partial \Psi}{\partial x} = U_z$ and $\frac{\partial \Psi}{\partial z} = -U_x$. Take linear combinations of the x and z derivatives of the resulting equations to eliminate the pressure. The Euler Lagrange equations become a fourth order equation:

$$\lambda \Delta \Psi = -\Delta^2 \Psi - Gr \frac{\partial^2 \Psi}{\partial x \partial z} \quad (20)$$

$$\frac{\partial \Psi}{\partial x} \Big|_{z=0,1} = 0 \quad (21)$$

$$\frac{\partial \Psi}{\partial z} \Big|_{z=0} = 0 \quad (22)$$

$$\frac{\partial^2 \Psi}{\partial z^2} \Big|_{z=1} = 0 \quad (23)$$

$$(24)$$

The system is periodic in x so we write Ψ in terms of its Fourier series. Let $k = \pi j/L$ be the wavenumber. We write $\Psi = \sum_k \hat{\Psi}_k e^{ikx}$. Then the equations become:

$$\lambda \left(\frac{\partial^2 \hat{\Psi}}{\partial z^2} - k^2 \hat{\Psi} \right) = - \left(\frac{\partial^4 \hat{\Psi}}{\partial z^4} - 2k^2 \frac{\partial^2 \hat{\Psi}}{\partial z^2} + k^4 \hat{\Psi} \right) - iGrk \frac{\partial \hat{\Psi}}{\partial z} \quad (25)$$

$$\hat{\Psi}|_{z=0,1} = 0 \quad (26)$$

$$\frac{\partial \hat{\Psi}}{\partial z}|_{z=0} = 0 \quad (27)$$

$$\frac{\partial^2 \hat{\Psi}}{\partial z^2}|_{z=1} = 0 \quad (28)$$

We have suppressed the k dependence of $\hat{\Psi}_k$. We search for the critical Grashoff number below which all of the eigenvalues λ are positive. This can be done with numerical methods. Each Fourier mode $\hat{\Psi}$ satisfies a fourth order ODE in z . We used finite differences to calculate the smallest eigenvalue corresponding to each k and determined the critical Grashoff number from this. After discretizing in z the differential equation becomes a generalized eigenvalue problem. We find that the critical Grashoff number is $Gr = 140$, at a value of $k = 3.1$.

2.2 Energy Stability for a Three Dimensional Stress Driven Flow

Consider the following non-dimensionalized equations for a three dimensional stress driven flow:

$$\frac{\partial u}{\partial t} + u \cdot \nabla u + \nabla p = \Delta u \quad (29)$$

$$\nabla \cdot u = 0 \quad (30)$$

$$u_z|_{z=0,1} = 0 \quad (31)$$

$$u_x|_{z=0} = 0 \quad (32)$$

$$u_y|_{z=0} = 0 \quad (33)$$

$$\frac{\partial u_x}{\partial z}|_{z=1} = Gr \quad (34)$$

$$\frac{\partial u_y}{\partial z}|_{z=1} = 0 \quad (35)$$

The domain is periodic in x and y . A steady solution to this equation is $u_x = Grz$. If we make the substitution $U = u + iGrz$ we can perform the same analysis as in the two dimensional case. The variational formulation is different and so is the set of Euler-Lagrange equations:

$$\lambda U_x = -\Delta U_x + \frac{Gr}{2} U_z + \frac{\partial q}{\partial x} \quad (36)$$

$$\lambda U_y = -\Delta U_y + \frac{\partial q}{\partial y} \quad (37)$$

$$\lambda U_z = -\Delta U_z + \frac{Gr}{2} U_x + \frac{\partial q}{\partial z} \quad (38)$$

If we assume that the eigenfunction for the lowest eigenvalue is not a function of x , we can introduce the stream function Ψ defined by $\frac{\partial \Psi}{\partial y} = U_z$ and $\frac{\partial \Psi}{\partial z} = -U_y$. The lowest frequency eigenmodes of shear driven flows tend to be Langmuir circulation flows that are independent of the streamwise direction. Our use of this assumption is justified by empirical observations within this field [6]. The Euler Lagrange equations become:

$$\lambda U_x = \frac{Gr}{2} \frac{\partial \Psi}{\partial y} - \Delta U_x \quad (39)$$

$$-\lambda \frac{\partial \Psi}{\partial z} = \frac{\partial q}{\partial y} + \Delta \frac{\partial \Psi}{\partial z} \quad (40)$$

$$\lambda \frac{\partial \Psi}{\partial y} = \frac{\partial q}{\partial z} - \Delta \frac{\partial \Psi}{\partial y} + \frac{Gr}{2} U_x \quad (41)$$

Taking the z derivative with respect to the second expression and the y derivative with respect to the first and third we can subtract the last two to eliminate the pressure:

$$\lambda \frac{\partial U_x}{\partial y} = \frac{Gr}{2} \frac{\partial^2 \Psi}{\partial y^2} - \Delta \frac{\partial U_x}{\partial y} \quad (42)$$

$$\lambda \Delta \Psi = -\Delta^2 \Psi + \frac{Gr}{2} \frac{\partial U_x}{\partial y} \quad (43)$$

We can write an ODE for the fourier modes of Ψ :

$$\lambda \imath k U_x = -\frac{Gr}{2} k^2 \hat{\Psi} - \imath k \left(\frac{\partial^2 U_x}{\partial z^2} - k^2 U_x \right) \quad (44)$$

$$\lambda \left(\frac{\partial^2 \hat{\Psi}}{\partial z^2} - k^2 \hat{\Psi} \right) = - \left(\frac{\partial^4 \hat{\Psi}}{\partial z^4} - 2k^2 \frac{\partial^2 \hat{\Psi}}{\partial z^2} + k^4 \hat{\Psi} \right) + \frac{Gr}{2} \imath k U_x \quad (45)$$

$$\hat{\Psi}|_{z=0,1} = 0 \quad (46)$$

$$\frac{\partial \hat{\Psi}}{\partial z}|_{z=0} = 0 \quad (47)$$

$$\frac{\partial^2 \hat{\Psi}}{\partial z^2}|_{z=1} = 0 \quad (48)$$

$$U_x|_{z=0} = 0 \quad (49)$$

$$\frac{\partial U_x}{\partial z}|_{z=1} = 0 \quad (50)$$

This is an ODE and we can discretize it and convert it into a generalized eigenproblem. We look for the critical Grashoff number Gr where the system first has a negative eigenvalue for some k . We find that $Gr_c = 51.7$ at $k = 2.1$. This is in agreement with work done by Tang, Caulfield, and Young [6].

2.3 Bounds on the Energy Dissipation in Two Dimensions

Define $\langle \cdot \rangle$ to be the space time average and $\bar{\cdot}$ to be the horizontal and time average. Begin by taking the dot product of the momentum equation with u . After averaging we find that:

$$\nu \langle |\nabla u|^2 \rangle = \frac{\tau}{h} \bar{u}_x(h) \quad (51)$$

Define the dissipation $\epsilon = \langle \nu |\nabla u|^2 \rangle$, Reynolds number of the flow to be $Re = \frac{\bar{u}_x(h)h}{\nu}$, and the friction coefficient to be $C_f = \frac{\epsilon h}{\bar{u}_x(h)^3}$. For the steady state solution $C_f(Re) = \frac{1}{Re}$. We want to determine a limit on how C_f might scale with the Reynolds number by proving bounds on the mean horizontal velocity at the $z = h$. Introduce a background field horizontal velocity $U(z)$ such that U satisfies the same boundary conditions as u . Consider the decomposition $u = \hat{i}U + \tilde{u}$. The following are the dimensionless equations for \tilde{u} :

$$\frac{\partial \tilde{u}}{\partial t} + \tilde{u} \cdot \nabla \tilde{u} + \nabla p + U \frac{\partial \tilde{u}}{\partial x} + \hat{i} \tilde{u}_z \frac{dU}{dz} = \Delta \tilde{u} + \hat{i} \frac{d^2 U}{dz^2} \quad (52)$$

$$\tilde{u}|_{z=0} = 0 \quad (53)$$

$$\tilde{u}_z|_{z=h} = 0 \quad (54)$$

$$\frac{\partial \tilde{u}_x}{\partial z}|_{z=h} = 0 \quad (55)$$

Take the dot product with \tilde{u} and compute the space time average:

$$0 = -\langle |\nabla \tilde{u}|^2 \rangle - \left\langle \frac{dU}{dz} \tilde{u}_x \tilde{u}_z \right\rangle + \left\langle \tilde{u}_x \frac{d^2 U}{dz^2} \right\rangle \quad (56)$$

The time derivative term vanishes because $\|\tilde{u}\|^2$ is bounded. If we integrate the last term by parts in the z direction the previous expression becomes:

$$0 = -\langle |\nabla \tilde{u}|^2 \rangle - \left\langle \frac{dU}{dz} \tilde{u}_x \tilde{u}_z \right\rangle + Gr \bar{u}_x(1) - \left\langle \frac{\partial \tilde{u}_x}{\partial z} \frac{dU}{dz} \right\rangle \quad (57)$$

$$= -\langle |\nabla \tilde{u}|^2 \rangle - \left\langle \frac{dU}{dz} \tilde{u}_x \tilde{u}_z \right\rangle + Gr(\bar{u}_x(1) - U(1)) - \left\langle \frac{\partial \tilde{u}_x}{\partial z} \frac{dU}{dz} \right\rangle \quad (58)$$

Substitute $\hat{i}U(z) + \tilde{u}(z) = \bar{u}(z)$ into $\langle |\nabla u|^2 \rangle$:

$$\frac{1}{2} \langle |\nabla u|^2 \rangle = \frac{1}{2} \langle |\nabla \tilde{u}|^2 \rangle + \frac{1}{2} \left\langle \frac{dU^2}{dz} \right\rangle + \left\langle \frac{\partial \tilde{u}_x}{\partial z} \frac{dU}{dz} \right\rangle \quad (59)$$

We take a linear combination of these expressions so as to eliminate the $\langle \frac{\partial \tilde{u}_x}{\partial z} \frac{dU}{dz} \rangle$ term:

$$\frac{1}{2} \langle |\nabla u|^2 \rangle = -\left\langle \frac{1}{2} |\nabla \tilde{u}|^2 + \frac{dU}{dz} \tilde{u}_x \tilde{u}_z \right\rangle + Gr(\bar{u}_x(1) - U(1)) + \frac{1}{2} \left\langle \frac{dU^2}{dz} \right\rangle \quad (60)$$

Now we define the quadratic form $Q_U(\tilde{u})$:

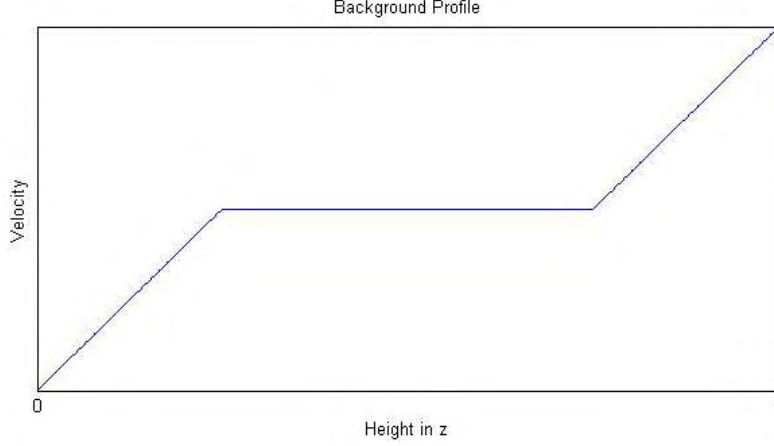
$$Q_U(\tilde{u}) = \left\langle \frac{1}{2} |\nabla \tilde{u}|^2 + \frac{dU}{dz} \tilde{u}_x \tilde{u}_z \right\rangle \quad (61)$$

Finally we combine this definition with the identity $\langle |\nabla u|^2 \rangle = Gr \bar{u}_x(1)$ to produce:

$$\bar{u}_x(1) = 2U(1) - \frac{1}{Gr} \left\langle \frac{dU^2}{dz} \right\rangle + \frac{2}{Gr} Q \quad (62)$$

If we choose U so that Q is positive definite we arrive at the following bound for $\bar{u}_x(1)$:

$$\bar{u}_x(1) \geq 2U(1) - \frac{1}{Gr} \left\langle \frac{dU^2}{dz} \right\rangle \quad (63)$$



Define the background horizontal velocity profile to be linear in z in horizontal boundary layers near the top and bottom boundary and constant elsewhere:

$$U(z) = Grz \quad 0 < z < \delta_1 \quad (64)$$

$$= Gr\delta_1 \quad \delta_1 < z < 1 - \delta_2 \quad (65)$$

$$= Gr(\delta_1 + \delta_2 + z - 1) \quad 1 - \delta_2 < z < 1 \quad (66)$$

Then $U(1) = Gr(\delta_1 + \delta_2)$ and $\bar{u}_x(1) \geq Gr(\delta_1 + \delta_2)$. We need to choose δ_1 and δ_2 to maximize the sum while keeping Q positive definite. We drop the accents and refer to the fluctuations away from the background as u .

$$Q = \frac{1}{2} \|\nabla u\|^2 + Gr \int_0^{\delta_1} u_x u_z dz dx + Gr \int_{1-\delta_2}^1 u_x u_z dx dz \quad (67)$$

We start with the second term on the right hand side:

$$\int \int_0^{\delta_1} u_x u_z dz dx = \int \int_0^{\delta_1} \int_0^z \left(\frac{\partial u_x}{\partial z'} u_z + u_x \frac{\partial u'_z}{\partial z'} \right) dz' dx dz \quad (68)$$

Now we use incompressibility to eliminate the $u_x \frac{\partial u'_z}{\partial z'} = -u_x \frac{\partial u_x}{\partial x} = -\frac{\partial \|u_x\|^2}{\partial x}$ term. Then we find:

$$\left| \int \int_0^{\delta_1} u_x u_z dz dx \right| = \left| \int_0^{\delta_1} \int_0^z \left(\frac{\partial u_x}{\partial z'} \int_0^{z'} \frac{\partial u''_z}{\partial z''} dz'' \right) dz' dx dz \right| \quad (69)$$

$$\leq \int \int_0^{\delta_1} \int_0^z \left| \frac{\partial u_x}{\partial z'} \right| \sqrt{z'} \left\| \frac{\partial u_z}{\partial z} \right\|_{\delta_1, z} dz' dx dz \quad (70)$$

$$\leq \int \int_0^{\delta_1} \frac{z}{\sqrt{2}} \left\| \frac{\partial u_x}{\partial z} \right\|_{\delta_1, z} \left\| \frac{\partial u_z}{\partial z} \right\|_{\delta_1, z} dx dz \quad (71)$$

$$\leq \frac{\delta_1^2}{4\sqrt{2}} \left(\frac{1}{C} \left\| \frac{\partial u_x}{\partial z} \right\|_{\delta_1}^2 + \frac{C}{2} \left\| \frac{\partial u_x}{\partial x} \right\|_{\delta_1}^2 + \frac{C}{2} \left\| \frac{\partial u_z}{\partial z} \right\|_{\delta_1}^2 \right) \quad (72)$$

Here $\|\cdot\|_{\delta_1} = \left(\int_0^{\delta_1} \cdot^2 dx dz\right)^{1/2}$ and $\|\cdot\|_{1-\delta_2} = \left(\int_{1-\delta_2}^1 \cdot^2 dx dz\right)^{1/2}$. Similarly define $\|\cdot\|_{\delta_1,z} = \left(\int_0^{\delta_1} (\cdot)^2 dz\right)^{1/2}$. If we choose $C = \frac{1}{\sqrt{2}}$ we find $\left|\int_0^{\delta_1} u_x u_z dz dx\right| \leq \frac{\delta_1^2}{16} \|\nabla u\|^2$. We can perform an identical analysis at the top boundary layer, and set $\delta_1 = \delta_2$ to prove that

$$Q \geq \frac{1}{2} \|\nabla u\|^2 - Gr \frac{\delta_1^2}{8} \|\nabla u\|_{\delta_1}^2 - Gr \frac{\delta_1^2}{8} \|\nabla u\|_{1-\delta_2}^2 \quad (73)$$

$$\geq \left(\frac{1}{2} - Gr \frac{\delta_1^2}{8}\right) \|\nabla u\|^2 \quad (74)$$

This is positive if $\delta_1 \leq 2Gr^{-1/2}$. This means that $\bar{u}_x(1) \geq 4Gr^{1/2}$. In terms of units $\bar{u}_x(h) = \bar{u}_x(1) \frac{\nu}{h} \geq 4\tau^{1/2}$. The friction coefficient $C_f = \frac{\tau}{\bar{u}(h)^2} \leq \frac{1}{16}$.

2.4 Bounds in Three Dimensions

We can use much of the same algebra to derive bounds for the three dimensional case. The equations are the same except that there is a y component in the velocity field. As a result we have the same $\bar{u}(h) \geq 2U(1) - \frac{1}{Gr} \left\langle \left(\frac{dU}{dz}\right)^2 \right\rangle$ as long as $Q \geq 0$. We pick the same background profile as in the two dimensional case, linear near each boundary and constant in the bulk.

We begin with the second term in Q and find that:

$$\left|\int \int \int_0^{\delta_1} u_x u_z dx dy dz\right| \leq \int \int \int_0^{\delta_1} z \left\| \frac{\partial u_x}{\partial z} \right\|_{\delta_1,z} \left\| \frac{\partial u_z}{\partial z} \right\|_{\delta_1,z} dx dy dz \quad (75)$$

$$\leq \frac{\delta_1^2}{2} \left(\frac{1}{2C} \int dx \int dy \left\| \frac{\partial u_x}{\partial z} \right\|_{\delta_1,z}^2 + \frac{C}{2} (1-a) \int dx \int dy \left\| \frac{\partial u_z}{\partial z} \right\|_{\delta_1,z}^2 + \frac{C}{2} a \int dx \int dy \left\| \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right\|_{\delta_1,y}^2 \right) \quad (76)$$

$$= \frac{\delta_1^2}{4} \left(\frac{1}{C} \left\| \frac{\partial u_x}{\partial z} \right\|_{\delta_1}^2 + C(1-a) \left\| \frac{\partial u_z}{\partial z} \right\|_{\delta_1}^2 + Ca \left\| \frac{\partial u_x}{\partial x} \right\|_{\delta_1}^2 + Ca \left\| \frac{\partial u_y}{\partial y} \right\|_{\delta_1}^2 + Ca \left\| \frac{\partial u_x}{\partial y} \right\|_{\delta_1}^2 + Ca \left\| \frac{\partial u_y}{\partial x} \right\|_{\delta_1}^2 \right) \quad (77)$$

Choose $a = \frac{1}{2}$ and $C = \sqrt{2}$ to find that $\left|Gr \int \int \int_0^{\delta_1} u_x u_z dx dy dz\right| \leq \frac{\delta_1^2}{4\sqrt{2}} \|\nabla u\|_{\delta_1}^2$. We make the transformation $z \mapsto 1-z$ to study the upper boundary layer:

$$\left|\int dx \int dy \int_0^{\delta_2} u_x u_z dz\right| \leq \int dx \int dy \int_0^{\delta_2} dz \frac{2}{\pi} \left(\int_{\delta_2}^1 \left(\frac{\partial u_x}{\partial z}\right)^2 dz'\right)^{1/2} \sqrt{z} \left\| \frac{\partial u_z}{\partial z} \right\|_{1-\delta_2,z} \quad (78)$$

$$\leq \frac{3\delta_2^{3/2}}{2\pi} \left(\frac{1}{C} \left\| \frac{\partial u_x}{\partial z} \right\|^2 + \frac{C}{2} \left\| \frac{\partial u_z}{\partial z} \right\|_{\delta_2}^2 + \frac{C}{2} \left\| \frac{\partial u_x}{\partial x} \right\|_{\delta_2}^2 + \frac{C}{2} \left\| \frac{\partial u_y}{\partial y} \right\|_{\delta_2}^2 \right) \quad (79)$$

Combining these two terms we find that:

$$|Q - \frac{1}{2} \|\nabla u\|^2| \leq Gr \frac{\delta_1^2}{4\sqrt{2}} \|\nabla u\|_{\delta_1}^2 + Gr \frac{3\delta_2^{3/2}}{2\pi} \left(\frac{1}{C} \left\| \frac{\partial u_x}{\partial z} \right\|^2 + \frac{C}{2} \|\nabla u\|_{\delta_2}^2 \right) \quad (80)$$

$$\leq Gr \left(\left[\frac{\delta_1^2}{4\sqrt{2}} + \frac{3\delta_2^{3/2}}{2\pi} \frac{1}{C} \right] \left\| \frac{\partial u_x}{\partial z} \right\|^2 + \frac{\delta_1^2}{4\sqrt{2}} \|\nabla u\|^2 \right) \quad (81)$$

$$= Gr \left(\left[\frac{\delta_1^2}{4\sqrt{2}} + \frac{3\sqrt{2}\delta_2^3}{2\pi^2\delta_1} \right] \left\| \frac{\partial u_x}{\partial z} \right\|^2 + \frac{\delta_1^2}{4\sqrt{2}} \|\nabla u\|^2 \right) \quad (82)$$

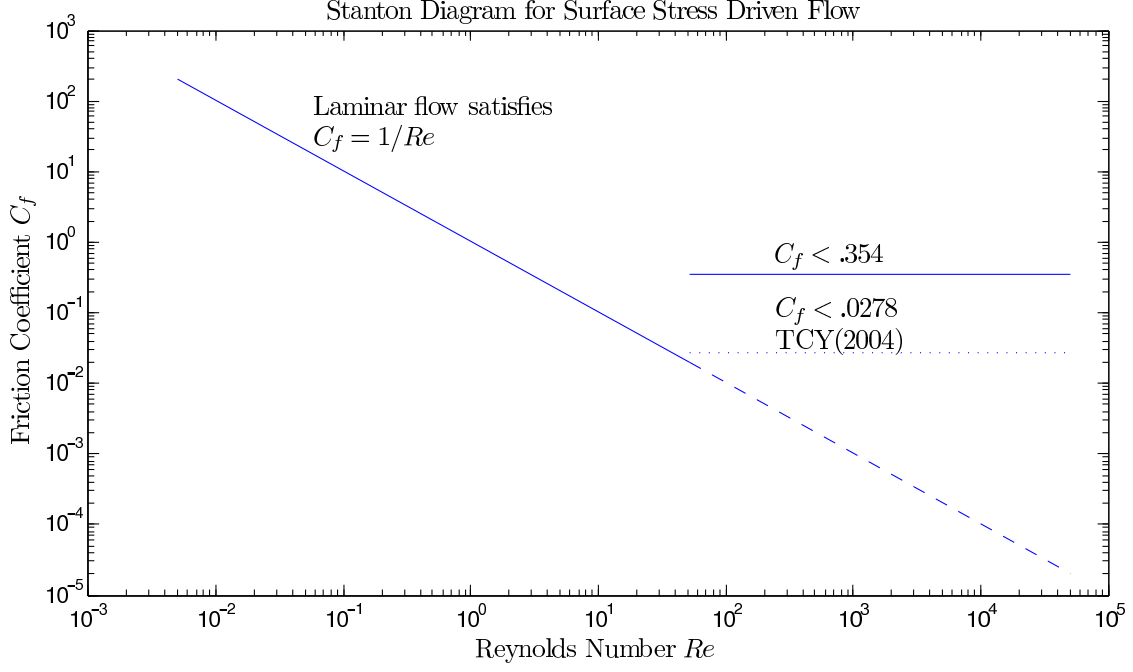


Figure 2: Stanton Diagram of the TCY bounds and the bounds proven here on C_f

In the last line we made the choice $C = \frac{\pi\delta_1^2}{3\sqrt{2}\delta_2^{3/2}}$. This shows that the ratio of the boundary layers $\frac{\delta_2}{\delta_1} \lesssim Gr^{-1/6}$. In order to produce a prefactor for the scaling of $\bar{u}(1)$ we take $\delta_2 = 0$. Then $\delta_1 = \sqrt{2\sqrt{2}Gr}^{-1/2}$ and $\bar{u}(h) \geq \sqrt{2\sqrt{2}\tau}^{1/2}$. The corresponding friction coefficient is $C_f \leq \frac{1}{2\sqrt{2}} = .354$. The bound proved by Tang, Caulfield, and Young[6] is $C_f < .0237$, which is stronger than this bound.

3 Infinite Prandtl Number Marangoni Convection

Consider the equations describing infinite prandtl number Marangoni convection:

$$\nabla p = \Delta u \quad (83)$$

$$\frac{\partial T}{\partial t} + u \cdot \nabla T = \Delta T \quad (84)$$

$$u|_{z=0} = 0 \quad (85)$$

$$T|_{z=0} = 0 \quad (86)$$

$$u_z|_{z=1} = 0 \quad (87)$$

$$\frac{\partial T}{\partial z}|_{z=1} = -1 \quad (88)$$

$$\frac{\partial u_x}{\partial z}|_{z=1} = -Ma \frac{\partial T}{\partial x}|_{z=1} \quad (89)$$

Take the divergence of the momentum equation:

$$\Delta p = \nabla \cdot \Delta u \quad (90)$$

$$= \Delta \nabla \cdot u \quad (91)$$

$$= 0 \quad (92)$$

If we take the Laplacian of the momentum equation we will find that u solves the biharmonic equation:

$$\Delta^2 u = \Delta \nabla p \quad (93)$$

$$= \nabla \Delta p \quad (94)$$

$$= 0 \quad (95)$$

We will be interested in the z component of the velocity for the background method. If we take the Fourier transform in the z direction it satisfies an ordinary differential equation:

$$\frac{d^4 u_z}{dz^4} - 2k^2 \frac{d^2 u_x}{dz^2} + k^4 u_x = 0 \quad (96)$$

$$u_z|_{z=0} = 0 \quad (97)$$

$$\frac{du_z}{dz}|_{z=0} = 0 \quad (98)$$

$$u_z|_{z=1} = 0 \quad (99)$$

$$\frac{d^2 u_z}{dz^2}|_{z=1} = k^2 Ma \theta(1) \quad (100)$$

This has an exact solution:

$$u_k = -Ma \theta_k(1) \frac{2 \sinh(k)}{\frac{2 \sinh(k) \cosh(k)}{k} - 2} (\sinh(kz) - kz \cosh(kz) + (k \coth(k) - 1) z \sinh(kz)) \quad (101)$$

Define $f_k(z) = \frac{u_k}{Ma \theta_k(1)}$.

We have plotted the modes in Figure 3. For large value of k the functions are concentrated near $z = 1$. For very large values of k the maximum value of f_k goes to zero. For small values of k the maximum of f_k is extremely small and the function is concentrated over the entire unit interval. For values of k near 3 and 4 the maximum is large and the concentration is also over a significant portion of the unit interval.

Set up the background method by making the substitution $T = \tau(z) + \theta$, where $\tau(0) = 0$ and $\frac{d\tau}{dz}(1) = -1$. Then the field θ satisfies the equation:

$$\frac{\partial \theta}{\partial t} + u_z \frac{d\tau}{dz} + u \cdot \nabla \theta = \Delta \theta + \frac{d^2 \tau}{dz^2} \quad (102)$$

$$\theta(z = 0) = 0 \quad (103)$$

$$\frac{\partial \theta}{\partial z}(1) = 0 \quad (104)$$

If we multiply the θ evolution equation by θ and integrate over space we get the expression:

$$\frac{1}{2} \frac{\partial \|\theta\|^2}{\partial t} = - \int dx \int dz (\theta u_z \tau' - \theta \tau'') - \|\nabla \theta\|^2 \quad (105)$$

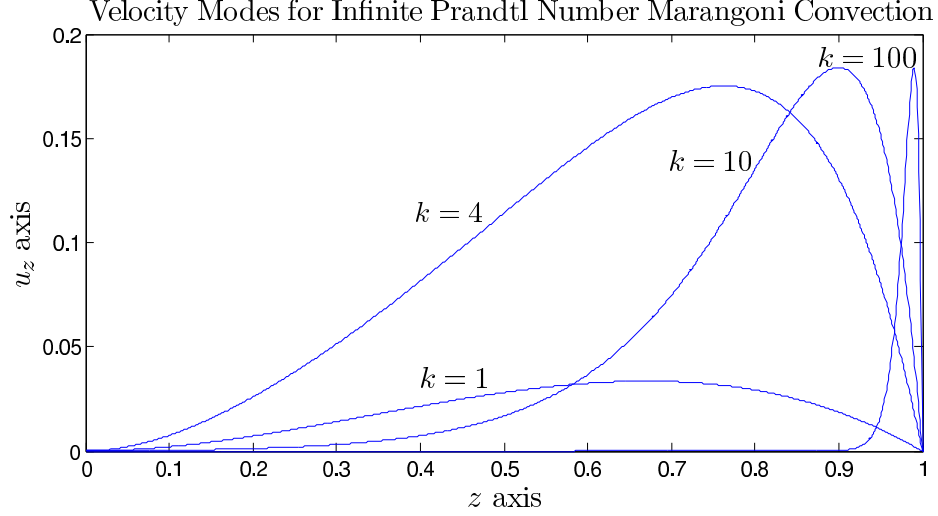


Figure 3: Plots of the functions f_k .

3.1 Energy Stability

Make the substitution $\tau = -z$ in the previous equation. The solution $\tau = -z$ will be stable when the quadratic form $Q = \|\nabla\theta\|^2 - \int u_z \theta dA$ is positive definite. Allow u_k to stand for the k th Fourier coefficient of u_z . Substitute the expression for the Fourier series of θ and u_z into Q :

$$Q = \sum_k \int_0^1 dz \left(-2\Re(\theta_k u_k) + \left(\frac{d\theta_k}{dz}\right)^2 + k^2 \theta_k^2 \right) dz \quad (106)$$

Then use the Cauchy-Schwarz inequality:

$$\left| \int_0^1 dz 2\Re(\theta_k u_k) \right| \leq Ma \theta_k(1) \|f_k\| \|\theta_k\| \quad (107)$$

$$\leq Ma \left\| \frac{d\theta_k}{dz} \right\| \|\theta_k\| \|f_k\| \quad (108)$$

$$\leq \frac{Ma}{2} \|f_k\| \left(k \left\| \frac{d\theta_k}{dz} \right\|^2 + \frac{1}{k} \|\theta_k\|^2 \right) \quad (109)$$

If for each k the quantity $\frac{Ma}{2k} \|f_k\|$ is less than 1 the form Q will be positive definite. Since we know the f_k we can write this condition explicitly:

$$\left(\int_0^1 dz \left[\frac{2 \sinh(k)}{2 \frac{\sinh(k)}{k} \cosh(k) - 2} (\sinh(kz) - kz \cosh(kz) + (k \coth(k) - 1) z \sinh(kz)) \right]^2 \right)^{1/2} \leq \frac{2k}{Ma} \quad (110)$$

This allows us to calculate the critical Marangoni number to be $Ma_c > 58.3$. The maximum occurs at $k = 2.4$. This estimate is slightly better than the estimate published Davis in

1969. Davis used variational methods and his bound is valid for all Prandtl numbers. The variational problem that he set up used natural boundary conditions to eliminate the terms that come from integrating by parts along the boundary. This weakening of the boundary condition is what makes it possible to improve the scaling using nonvariational methods.

3.2 Upper Bounds on the Nusselt Number

We define the Nusselt number to be the ratio of heat transport to heat transport due to conduction, so that it measures the strength of convection. Conduction requires a temperature gradient, so one would expect that when convection is strong the fixed heat flux is maintained with a minimal temperature gradient. The result is that the temperature at the top becomes low. Therefore we define the Nusselt number as $-\frac{1}{T(z=1)}$. If we multiply the time evolution equation for T by T and average we find that:

$$\frac{1}{Nu} = \|\nabla T\|^2 \quad (111)$$

Take the energy evolution equation for $\|\theta\|^2$ and time average it:

$$0 = - \int dx \int dz (\theta u_z \tau' - \theta \tau'') - \|\nabla \theta\|^2 \quad (112)$$

Now use the expression $\|\nabla \theta\|^2 = \|\nabla T\|^2 - \|\tau'\|^2 - 2 \int dx dz \tau' \frac{\partial \theta}{\partial z}$ to write:

$$0 = -2 \int dx \int dz (\theta u_z \tau' - \theta \tau'') - \|\nabla \theta\|^2 - \|\nabla T\|^2 + 2 \int dx dz \tau' \frac{\partial \theta}{\partial z} + \|\tau'\|^2 \quad (113)$$

$$0 = -2 \int dx \int dz (\theta u_z \tau') - \|\nabla T\|^2 - 2\bar{\theta}(1) - \|\nabla \theta\|^2 + \|\tau'\|^2 \quad (114)$$

$$0 = -2 \int dx \int dz (\theta u_z) - 2\bar{T}(1) + 2\bar{\tau}(1) - \|\nabla \theta\|^2 + \|\tau'\|^2 + \|\nabla T\|^2 \quad (115)$$

$$-\frac{1}{Nu} = -2 \int dx \int dz (\theta u_z \tau') - \|\nabla \theta\|^2 + 2\bar{\tau}(1) + \|\tau'\|^2 \quad (116)$$

If the functional $Q = 2 \int dx \int dz (\theta u_z \tau') + \|\nabla \theta\|^2$ is always positive we can prove an upper bound on the Nusselt number:

$$\frac{1}{Nu} \geq -2\bar{\tau}(1) - \int dz \tau'^2 \quad (117)$$

We choose the derivative of the background profile τ to be equal to -1 in a layer of width δ_1 near the bottom and δ_2 near the top. We choose τ' to be constant in the bulk:

$$\tau(z) = -z \quad 0 < z < \delta_1 \quad (118)$$

$$= -\delta_1 \quad \delta_1 < z < 1 - \delta_2 \quad (119)$$

$$= -\delta_1 - z + 1 - \delta_2 \quad 1 - \delta_2 < z < 1 \quad (120)$$

Then $\frac{1}{Nu} \geq \delta_1 + \delta_2$ as long as $\int dx \int dz (2\theta u_z) + \|\nabla \theta\|^2$ is positive. Write this in terms of the Fourier decomposition of θ and u_z :

$$0 \leq \sum_k \left[2 \int_0^{\delta_1} dz \Re(\theta_k u_k) + 2 \int_{1-\delta_2}^1 dz \Re(\theta_k u_k) + \left\| \frac{d\theta_k}{dz} \right\|^2 + k^2 \|\theta_k\|^2 \right] \quad (121)$$

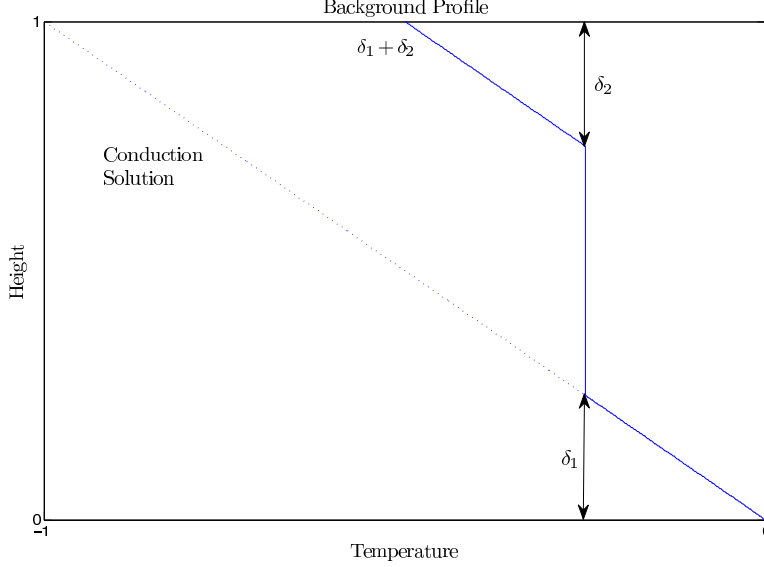


Figure 4: Background profile for Marangoni convection.

If for all k the quantity $2 \int_0^{\delta_1} dz \Re(\theta_k u_k) + 2 \int_{1-\delta_2}^1 dz \Re(\theta_k u_k) + \|\frac{d\theta_k}{dz}\|^2 + k^2 \|\theta_k\|^2 > 0$ the form Q will be positive definite.

We begin the analysis with the lower boundary layer. We start by replacing u_k with $Ma\theta_k(1)f_k$. Then we use the fundamental theorem of calculus:

$$\left| \int_0^{\delta_1} f_k \theta_k(1) \theta_k dz \right| \leq \int_0^{\delta_1} |f_k \theta_k| dz \left\| \frac{d\theta_k}{dz} \right\| \quad (122)$$

$$\leq \int_0^{\delta_1} \sqrt{z} |f_k| \left\| \frac{d\theta_k}{dz} \right\|^2 dz \quad (123)$$

Let $F(z) = \sup_k f_k(z)$. Then we need to evaluate $\int_0^{\delta_1} \sqrt{z} F(z) dz$ to determine the scaling. As z goes to zero each f_k scales as $c_k z^2$ so there is reason to believe that F might have the same behavior. We verified numerically that for extremely small z the value of k at which the maximum over f_k is realized has a lower bound, implying that there is some c such that $F(z) < cz^2$. We show this in Figure 3.2. This means we can choose $\delta_1^{7/2} = O(Ma^{-1})$, or that $Nu < O(Ma^{2/7})$. We can calculate the prefactor numerically. We set $\delta_2 = 0$. We find that $Nu < .84 Ma^{2/7}$.

Boeck and Thess have used numerical methods to find solutions of the infinite Prandtl number Marangoni problem. They analyzed the scaling of the Nusselt number of their numerical solutions with respect to the Marangoni number and found that $Nu = .446 Ma^{.238}$. They also theorized that $Nu \sim Ma^{2/9}$ [1]. We compare their data to our bound in Figure 3.2.

3.3 Heat Transport with finite Prandtl Number

We consider the problem of Marangoni Convection in a system with finite Prandtl number. The system is described by the equations:

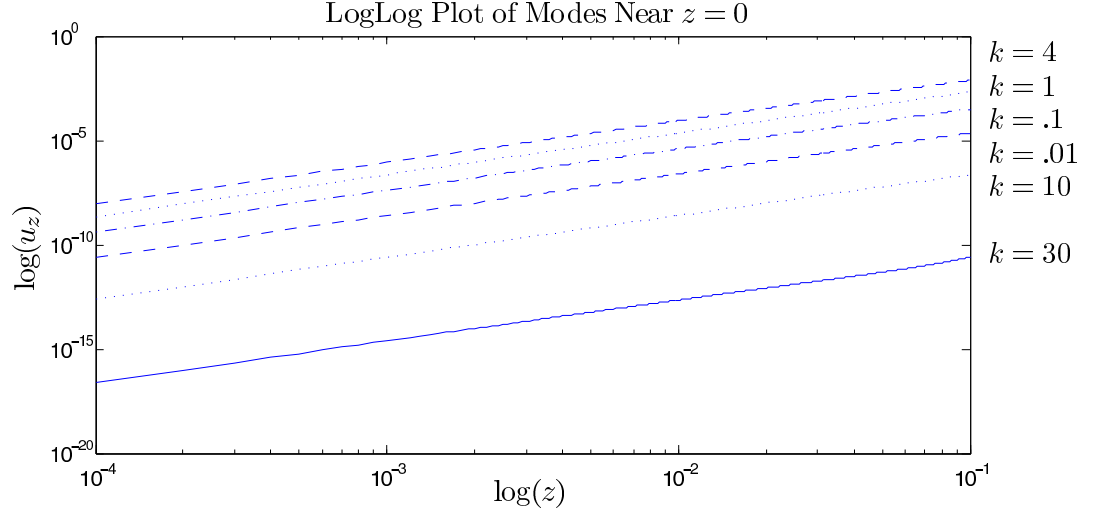


Figure 5: Scaling of the functions f_k for different values of k near $z = 0$.

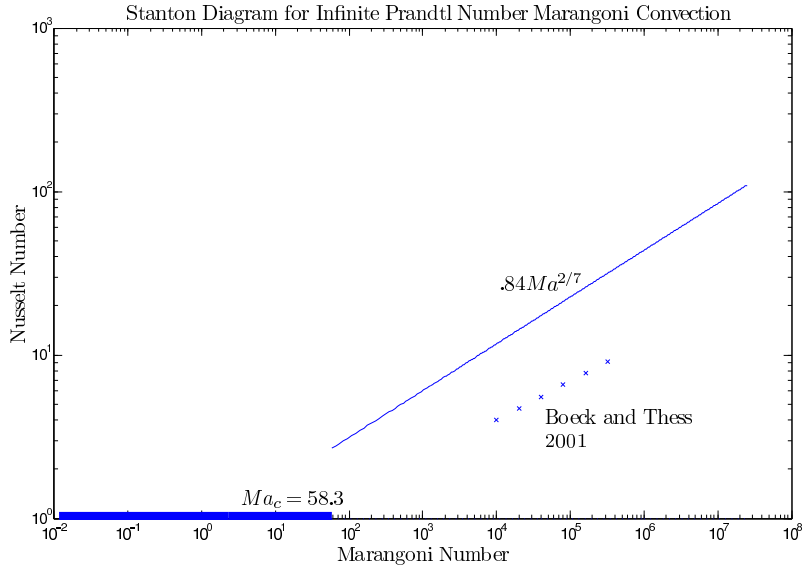


Figure 6: Plot of Nusselt and Marangoni number of numerical data compared to the rigorous upper bound and the critical Marangoni number

$$P^{-1}(\frac{\partial u}{\partial t} + u \cdot \nabla u) = -\nabla p + \Delta u \quad (124)$$

$$\nabla \cdot u = 0 \quad (125)$$

$$\frac{\partial T}{\partial t} + u \cdot \nabla T = \Delta T \quad (126)$$

$$u(z=0) = \vec{0} \quad (127)$$

$$T(z=0) = 1 \quad (128)$$

$$u_3(z=1) = 0 \quad (129)$$

$$\frac{\partial T}{\partial z}(z=1) = -1 \quad (130)$$

$$\frac{\partial u_1}{\partial z} = -Ma \frac{\partial T}{\partial x}(z=1) \quad (131)$$

$$\frac{\partial \theta}{\partial z}(z=1) = 0 \quad (132)$$

The Nusselt number of this flow is defined as $Nu = -\frac{1}{T(1)}$. Multiply the equation for the temperature by T and average it over the volume. All the terms on the left hand side vanish, the first term because the norm of temperature is bounded and the advection term because there is no flux across the upper and lower boundaries. Therefore we find that:

$$0 = \int_V T \Delta T dV \quad (133)$$

$$= -\|\nabla T\|^2 + \int_{\Gamma V} T \frac{\partial T}{\partial z} dA \quad (134)$$

$$= -\int_{z=1} T dA - \|\nabla T\|^2 \quad (135)$$

This means that $Nu = \frac{1}{\|\nabla T\|^2}$. Introduce a background profile $\tau(z)$. We choose τ to satisfy a homogeneous Dirichlet condition at $z=0$ and the Neumann condition $\frac{\partial \tau}{\partial z} = -1$ at $z=1$. Then we can decompose the temperature $T = \tau(z) + \theta(x, z, t)$ where $\theta = 0$ at $z=0$ and $\frac{\partial \theta}{\partial z} = 0$ at $z=1$. For each τ there is an equation for θ :

$$\frac{\partial \theta}{\partial t} + u \cdot \nabla \theta + u_z \frac{\partial \tau}{\partial z} = \Delta \theta + \frac{\partial^2 \tau}{\partial z^2} \quad (136)$$

Multiply this equation by θ and take the average over space and time. Again the first two terms on the left hand side vanish:

$$0 = -\|\nabla \theta\|^2 - \bar{\theta}(1) - \langle \theta u_z \tau' \rangle - \langle \frac{\partial \theta}{\partial z} \rangle \quad (137)$$

Now substitute τ and θ for T into $\|\nabla T\|^2 = \frac{1}{Nu}$:

$$\frac{1}{Nu} = \|\nabla \theta\|^2 + \|\nabla \tau\|^2 + 2 \int \nabla \theta \cdot \nabla \tau dV \quad (138)$$

$$= -\|\nabla \theta\|^2 - 2\bar{\theta}(1) - 2\langle \theta u_z \tau' \rangle + \int \tau'^2 dv \quad (139)$$

In the last line we added two times the identity in the previous equation. In order to incorporate the Marangoni condition we need the momentum equation. We take the dot product of the momentum equation with u and calculate the space time average:

$$0 = -\|\nabla u\|^2 + \int_{z=1} u_x \frac{\partial u_x}{\partial z} dA \quad (140)$$

$$= -\|\nabla u\|^2 - Ma \int_{z=1} u_x \frac{\partial \theta}{\partial x} dA \quad (141)$$

$$= -\|\nabla u\|^2 - Ma \int_V \frac{\partial}{\partial z} (u_x \frac{\partial \theta}{\partial x}) dV \quad (142)$$

Consider linear combinations of this expression and the equation for the Nusselt number:

$$\frac{1}{Nu} = -\|\nabla \theta\|^2 - 2\bar{\theta}(1) - 2\langle \theta u_z \tau' \rangle + \int_V \tau'^2 dv + C(\|\nabla u\|^2 + Ma \int_V \frac{\partial}{\partial z} (u_x \frac{\partial \theta}{\partial x}) dV) \quad (143)$$

$$= -\|\nabla \theta\|^2 - 2(\bar{T}(1) - \bar{\tau}(1)) - 2\langle \theta u_z \tau' \rangle + \int_V \tau'^2 dv + C(\|\nabla u\|^2 + Ma \int_V \frac{\partial}{\partial z} (u_x \frac{\partial \theta}{\partial x}) dV) \quad (144)$$

$$= -2\bar{\tau}(1) - \langle \tau'^2 \rangle + Q(\theta, \tau, u) \quad (145)$$

In the second to last line we used the fact that $\theta = T - \tau$. In the last line we used the fact that $Nu = -\frac{1}{T(1)}$ and defined the quadratic form Q :

$$Q(\theta, \tau, u) = \|\nabla \theta\|^2 + 2\langle \theta u_z \tau' \rangle + C(\|\nabla u\|^2 + Ma \int_V \frac{\partial}{\partial z} (u_x \frac{\partial \theta}{\partial x}) dV) \quad (146)$$

Suppose that $Q > 0$ for all u, θ , for some given choice of τ . Then $\frac{1}{Nu} < -2\bar{\tau}(1) - \langle \tau'^2 \rangle$. The bound on the Nusselt number will only depend on the choice of τ . Therefore the task will be to choose a τ that maximizes the bound on $\frac{1}{Nu}$ subject to the constraint that Q is positive. The background τ only affects Q through the term $\langle \theta u_z \tau' \rangle$, which can be negative. Therefore it would make sense to make this term as small as possible. Define τ such that τ' is zero everywhere except near the upper and lower boundaries.

$$\tau = -z \quad z < \delta_1 \quad (147)$$

$$\tau = -\delta_1 \quad \delta_1 < z < 1 - \delta_2 \quad (148)$$

$$\tau = -\delta_1 - \delta_2 + 1 - z \quad 1 - \delta_2 < z < 1 \quad (149)$$

With this choice of τ the expression $\langle \theta u_z \tau' \rangle$ reduces to the inner product of u_z and θ over regions of width δ_1 and δ_2 near the boundaries. The τ also satisfies the boundary conditions. We treat the top and bottom layer separately, starting with the bottom layer:

$$\int_0^{\delta_1} \theta u_z dV \leq \int_0^{\delta_1} \left(\int_0^z \left| \frac{\partial}{\partial z'} \theta dz' \right| \right) \left(\int_0^z \left| \frac{\partial}{\partial z''} u_z dz'' \right| \right) dx dz \quad (150)$$

$$\leq \int_0^{\delta_1} z \left\| \frac{\partial \theta}{\partial z} \right\| \left\| \frac{\partial u_z}{\partial z} \right\| dx dz \quad (151)$$

$$= \frac{\delta_1^2}{2} \left\| \frac{\partial \theta}{\partial z} \right\| \left\| \frac{\partial u_z}{\partial z} \right\| \quad (152)$$

It is more difficult to deal with the upper layer because θ is not zero on the upper boundary. As a result we have to use the Poincare inequality instead of the method we just used on the lower layer. We can choose $\delta_2 = 0$ and not effect the scaling.

The term involving the Marangoni number must also be controlled by the norms of the derivatives of u and θ .

$$Ma \int \frac{\partial}{\partial z} \left(u_x \frac{\partial \theta}{\partial x} \right) = Ma \int \left(\frac{\partial u_x}{\partial z} \frac{\partial \theta}{\partial x} - \frac{\partial u_x}{\partial x} \frac{\partial \theta}{\partial z} \right) \quad (153)$$

$$\leq \left\| \frac{\partial u_x}{\partial z} \right\| \left\| \frac{\partial \theta}{\partial x} \right\| + \left\| \frac{\partial u_x}{\partial x} \right\| \left\| \frac{\partial \theta}{\partial z} \right\| \quad (154)$$

Here we have used the fact that u is incompressible. Now we have bound on Q in terms of first derivatives of u and θ . We are only interested in the scaling and do not make an effort here to derive an optimal bound. If we take $C = 1$ we find that Q is positive when the following equation for λ has no negative solutions:

$$(1 - \lambda)^4 - \left(\frac{\delta^4}{4} + \frac{Ma\delta^2}{2} \right) (1 - \lambda)^2 = 0 \quad (155)$$

This is satisfied as long as $\frac{\delta^4}{4} + \frac{Ma\delta^2}{2} < 1$. This suggests that for large Ma $\delta \lesssim Ma^{-1/2}$. This suggests that $Nu \lesssim Ma^{1/2}$.

The scaling $Nu \lesssim Ma^{1/2}$ is considerably weaker than the infinite Prandtl number case. On the other hand, simulations by Boeck and Thess do not indicate that the actual scaling is weaker for the finite Prandtl number case. The reason that this scaling is so much worse is because we were unable to take advantage of the relationship between the temperature and velocity fields. It was possible to achieve the 2/7th bound because we knew what the Fourier modes of the velocity field were. However the relationship is much more complicated in the finite Prandtl number case and we were unable to use it. In general, improvements of the bounds produced by the background method in Raleigh-Benard convection were achieved by using the relationship [2], so it seems reasonable to expect that this bound could be improved by such considerations.

4 Conclusion

We presented rigorous upper bounds for the friction coefficient in stress driven shear flow and also derived the critical Grashoff number for the stability of the laminar flow solution. Tang, Caulfield, and Young first used the background method to analyze this problem. We were able to prove an analytic bound on the friction coefficient using the full stress boundary condition. We improved the lower bound on the critical Marangoni number for nonlinear stability in infinite dimensional Prandtl number Marangoni convection. We also proved that $Nu < .84Ma^{2/7}$ for the same problem. In finite Prandtl number Marangoni convection we were unable to do any better than $Nu \sim Ma^{1/2}$.

References

- [1] T. BOECK AND A. THESS, *Power-law scaling in benard-marangoni convection at large prandtl numbers*, Phys. Rev. E, 64 (2001).

- [2] F. O. C. DOERING AND M. REZNIKOFF, *Bounds on vertical heat transport for infinite prandtl number rayleigh-benard convection*, Flu. Mech., 560 (2006), pp. 229–241.
- [3] S. DAVIS, *Buoyancy surface tension instability by the method of energy*, Flu. Mech., 39 (1969), pp. 347–359.
- [4] C. DOERING AND P. CONSTANTIN, *Energy dissipation in shear driven turbulence*, Phys. Rev. A, 69 (1992), pp. 1648–1651.
- [5] J. PEARSON, *On convection cells induced by surface tension*, Flu. Mech., 4 (1958), pp. 489–500.
- [6] C. C. W. TANG AND W. YOUNG, *Bounds on dissipation in stress driven flow*, Flu. Mech., 510 (2004), p. 333.

Continual skipping - getting a kick from water

Ian Hewitt

1 Introduction

Attempting to ‘skip’ stones at the beach is familiar to many - throwing the stone into the water surface in such a way that it bounces several times before eventually sinking. This project investigates whether, by towing an object along at a sustained horizontal velocity, it can *continually* skip along the surface, and looks at the fluid dynamics which might cause this to happen.

We conduct experiments in which we tow an inclined plane (which we refer to here as a paddle) through the water at different speeds and observe whether it glides steadily through the water or whether it oscillates and skips clear of the surface. We then attempt to explain the observations with some simple models for the height of the paddle.

We find that a light paddle will continually skip provided the water speed is large enough, and the size of the bounces increases with the water speed above this threshold. We can study how this depends on the mass of the paddle, and also its width.

The steady planing state, at least when treated as two dimensional, is a classical problem [4, 16] and also bears some similarity to steady flow beneath a sluice gate [18, 1] or a surfboard [17]. However the stability of this state has not received much attention and allowing for time dependence significantly alters the problem.

A related body of research looks at ‘water entry’ of an object, the majority of which has military applications; particularly the impact of torpedoes and bombs on the surface and the forces on a seaplane’s floats during landing [19, 20, 8]. The main questions in this case are whether the force exerted by the water on the entering object is large enough to cause it to ricochet off the surface (usually not desired for torpedoes, but famously put to effect by Barnes Willis’s ‘bouncing bomb’), or whether the seaplane will maintain its stability as a result of the force experienced on the floats.

Mathematical models of the response of both shallow and deep water due to an impacting object have been developed [9, 20, 7], looking at the wetted region of the object, the high speed jets which are ejected and the force on the object; taking into account all the details of the water flow makes the results reasonably complicated. A simpler description of the force on the object was used recently to describe experiments on skipping stones [3, 13], and we use this as a starting point for a description of our experiments. It will be seen that this model is too simplistic and we must add some extra physics - namely the ‘splash-up’ of water ahead of the object - to explain the observed skipping.

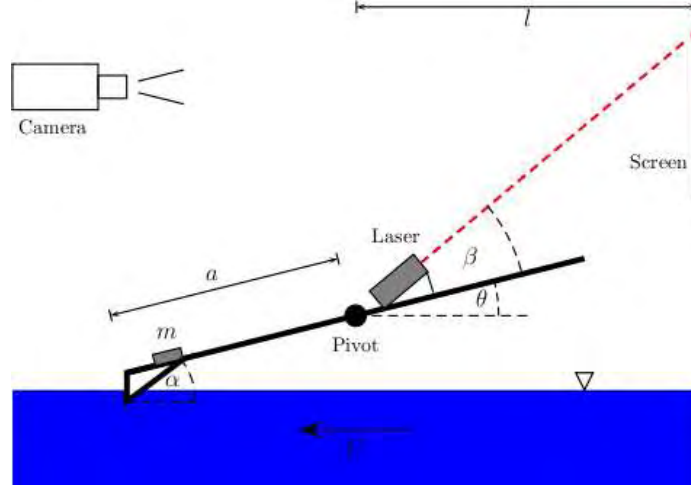


Figure 1: Experimental setup. A counterbalanced wooden arm of length $2a$ is pivoted at a fixed point, with a paddle attached to one end and a laser attached to the other. The paddle side is weighed down by a mass m . The paddle is usually 7 cm wide, although some experiments used a wider 10 cm paddle. The laser shines a beam at an angle β onto a screen where its position is captured by a video camera at 30 frames per second.

2 Experiments

The setup for the experiments is shown in figure 1. We performed most of the experiments in a large (2 m diameter) rotating tank with the paddle held by an arm onto a fixed pivot. By changing the angular velocity of the rotating tank we could adjust the water speed beneath the paddle and arm, which were aligned parallel to the water flow near the edge of the tank.

The arm holding the paddle was counterbalanced on the other side so that the weight of the paddle pulling it down could be made small. By adding additional masses on top of the paddle we could then increase its weight. Attached to the arm was a laser pointer which shined its beam upwards onto a screen (see figure 1). The screen could be recorded with a video camera, and the position of the laser obtained from analysing the images. We used a standard 30 frames per second analog recorder, so could extract the position of the laser every $1/30$ of a second. The position of the laser can then be converted into the vertical position of the paddle so that we obtain the trajectory of the paddle sampled at 30 Hz.

The water flowing beneath the paddle was about 6 cm deep, but one of the drawbacks of using a rotating tank was that this depth depended on the flow speed. The surface of the water is not flat - it is parabolic in shape - and becomes more and more inclined as the angular velocity of the tank increases. In order to ensure that the paddle bottom was as near as possible to parallel to the surface (since the surface is parabolic and the paddle straight, this is not possible exactly) the arm holding the pivot had to be angled differently for different water speeds.

The tank could be rotated at up to 2 rad s^{-1} , and the paddle was positioned about 20 cm from the edge of the tank so that the highest possible water speed was 1.76 m s^{-1} .

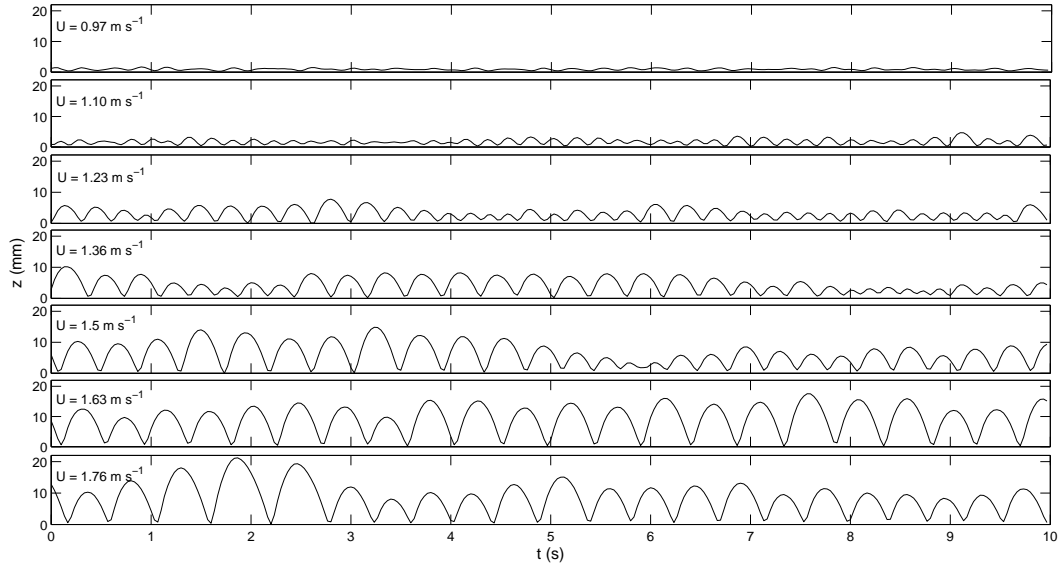


Figure 2: 10 s segments of time series for increasing water speed U and mass $m = 3.4$ g. This is the vertical displacement of the paddle, but the origin in each case is arbitrarily chosen to be the minimum position, and is therefore not quite the same for each speed. The height of the undisturbed water surface is not known exactly but will be between 0 – 5 mm above the minimum. At lower speeds the paddle sits in a steady planing state without oscillating.

This speed is high enough to cause a light paddle to continually skip along the surface - indeed there seemed to be no steady planing state in this case and independently of how the paddle was dropped or placed into the water it would evolve into a bouncing trajectory which had a characteristic amplitude and frequency. Whilst each bounce was somewhat irregular - they were not all the same amplitude - the amplitude of the bouncing mode over a long enough time series was quite well defined. The trajectory was recorded in 45 second time series which were begun long after the paddle was placed in the water - these trajectories are therefore indicative of the ‘steady’ bouncing behaviour. Examples of the trajectories obtained are shown in figures 2 and 3. The origin for each of these trajectories is not the same - due to the experimental setup it was not possible to accurately determine where on the screen corresponds to the undisturbed water surface, so the displacements measured can only be regarded as *relative* displacements. It is also not easy to determine when the paddle leaves the water surface - we were only able to assess this visually. It appears that small oscillations are possible with the paddle never leaving the surface - for instance below speed $U = 1.1$ m s⁻¹ in figure 2 and above mass $m = 17.3$ g in figure 3.

At high water speeds the rotation period of the tank is only just over 3 seconds, which does not give long for the waves created in the wake of the paddle to be carried round into its path. We would like the water ahead of the paddle to be as smooth as possible, since we are looking for natural oscillations rather than setting up a wave pattern around the tank (unlike [15], for example), so ideally we would like to remove the wake entirely. This

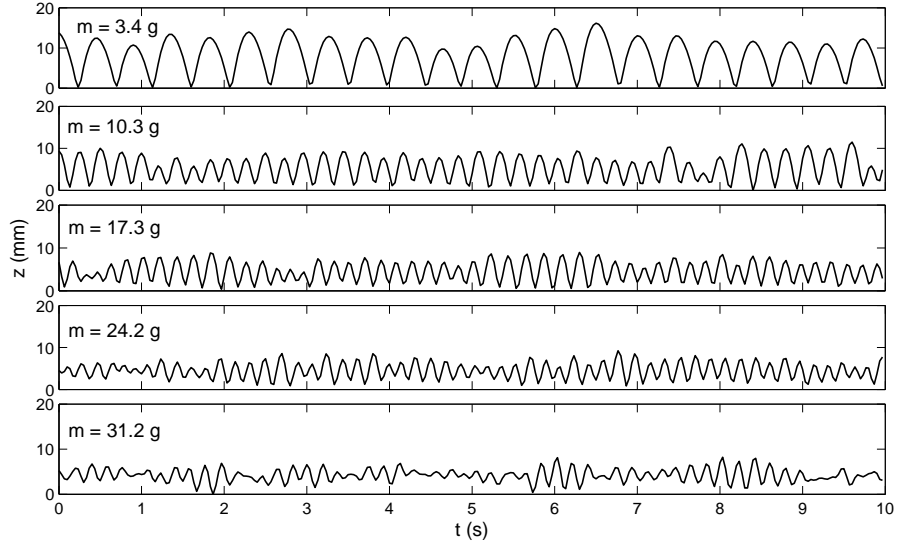


Figure 3: 10 s segments of time series for 5 different masses, with water speed $U = 1.5 \text{ m s}^{-1}$. This is the vertical displacement of the paddle, but the origin in each case is arbitrarily chosen to be the minimum position, and is therefore not quite the same for each mass. The height of the undisturbed water surface is not known exactly but will be around 5 mm above the minimum.

is another drawback of the rotating tank setup, and we attempted to minimise the wave activity ahead of the paddle by padding the sides of the tank with netting, to damp the waves which reflect off the walls. At large water speeds the center of the tank was also dry so we similarly padded the inner ‘shoreline’ where the waves would otherwise break and reflect. This had some effect in smoothing out the water, though we cannot claim to have removed the waves entirely. However, we do not believe that the presence of waves ahead of the paddle is what causes the oscillations - although it may lead to a more noisy signal than we could expect if the water surface were really smooth. By dropping the paddle into smooth water it was observed that the regular amplitude of bounces was set up *before* the first wake had travelled round the tank, suggesting that excitation by the waves is not the driving mechanism.

The main parameters we varied were the speed of the water U and the mass on the paddle m , as shown in figures 2 and 3. The power spectrum of these trajectories has a peak at frequencies close to the evident natural frequency of oscillations, but it is not very well defined, and the amplitude of the peak is not a good measure of the amplitude of oscillations. This might be expected since when the paddle leaves the water it ought to follow a parabolic ballistic trajectory, whereas in the water it may be more sinusoidal; the combination cannot be expected to decompose clearly in Fourier coefficients.

Instead we analyse the data by isolating individual bounces or oscillations and calculating their amplitude and period. Each 45 second time series contains over 100 bounces so we can look at the mean and standard deviation; these are shown in figures 4 and 5. The standard deviation gives a measure of how variable the amplitude and period are and there-

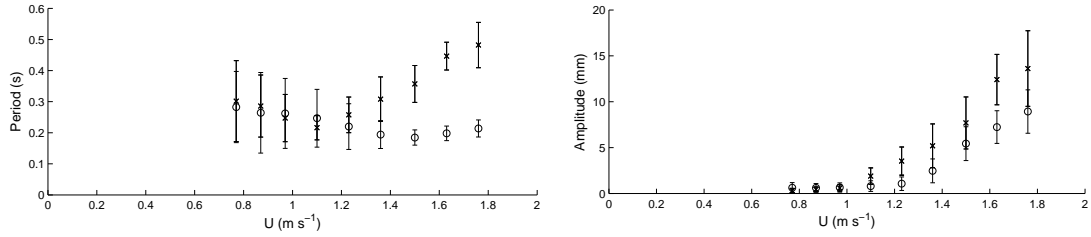


Figure 4: Period and amplitude for a paddle of mass 3.4 g (crosses) and mass 17.3 g (circles), with different water speeds U . These are average periods and amplitudes of oscillations over a 45 s time interval. The error bars are one standard deviation either side of the mean, and therefore give an indication of the variability rather than the error.

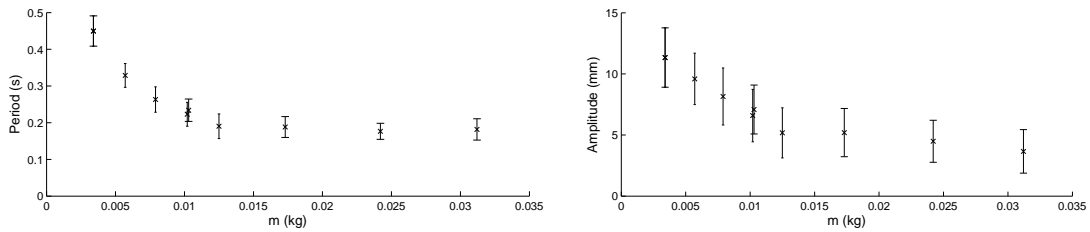


Figure 5: Period and amplitude for water speed $U = 1.5 \text{ m s}^{-1}$, with different paddle masses m . This shows the average period of oscillation and amplitude, with the error bars showing one standard deviation either side of the mean. These are taken from the amplitudes and periods of bounces within a 45 s time interval.

fore how well defined the bouncing mode is. At slow speeds, the trajectories are seen to be considerably more noisy and irregular and this is evident in the larger standard deviation for the period.

To summarise what these data show, for a given paddle mass there seems to be a threshold water speed above which oscillations start to occur (figure 4), and above this, the amplitude of the oscillations increases, apparently linearly, with the water speed. This is suggestive of some sort of bifurcation, and we also see that the threshold seems to be higher for a heavier paddle. When the bounces have large amplitude, there is a simultaneous increase in the period, as we might expect for ballistic trajectories.

For a constant water speed (figure 3), it appears that the lighter mass undergoes larger oscillations with a longer period and as the mass is increased the amplitude and period decrease, with the period levelling off towards a roughly constant 0.2 s.

These are some of the trends which we hope to be able to explain, at least qualitatively, with a simple model.

We also conducted some experiments with a wider paddle of width 10 cm to examine the effect of the paddle's width. For the same mass, the wider paddle undergoes larger amplitude bounces and appears to require a lower threshold water speed to start oscillating (figures 6 and 7). This dependence on the width of the paddle is another feature which we

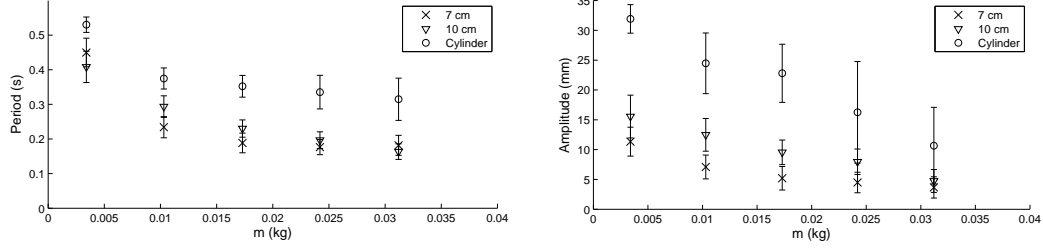


Figure 6: Period and amplitude for water speed 3.5 m s^{-1} , for 5 different masses with 7 cm and 10 cm wide paddles and with a cylinder shaped paddle. These are average periods and amplitudes of oscillations over a 45 s time interval. The error bars are one standard deviation either side of the mean.

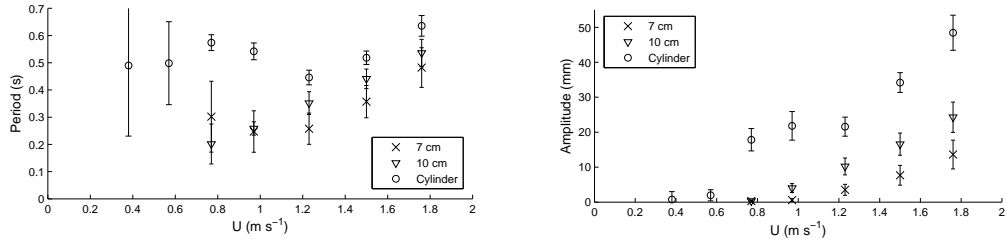


Figure 7: Period and amplitude for a paddle of mass 3.4 g for different water speeds with 7 cm and 10 cm wide paddles and with a cylinder shaped paddle. These are average periods and amplitudes of oscillations over a 45 s time interval. The error bars are one standard deviation either side of the mean.

hope the model will capture.

Though we will not discuss them further we also did some tests with a cylindrical paddle (lengthways across the stream). This oscillated very regularly and with large amplitude and the threshold to begin oscillating was at a considerably lower water speed. The data are shown in figures 6 and 7 for completeness.

3 Modeling preliminaries

The vertical position of the paddle is governed by the angle θ between the arm and the horizontal, so the equation of motion is really conservation of angular momentum about the pivot. The equation is

$$\mathcal{I}\ddot{\theta} = -mga \cos \theta + Fa \cos \theta - \mu a^2 \dot{\theta}, \quad (1)$$

in which \mathcal{I} is the moment of inertia, a is the length of the arm, m is the excess weight on the paddle side, g is the gravitational acceleration, and F is the vertical force of the water on the paddle. μ is a friction coefficient which describes the damping by friction in the hinge and air resistance. The vertical displacement of the paddle is $z = a \sin \theta$ and since θ

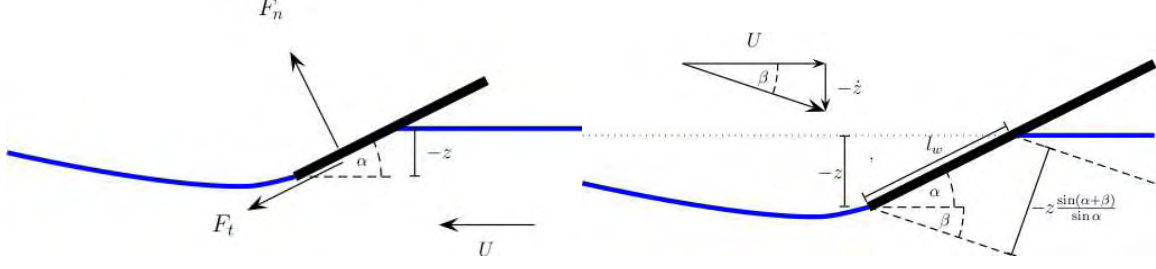


Figure 8: (a) Forces on a paddle inclined at an angle α to the horizontal water surface, in a stream moving at speed U . z , measured vertically upwards, is the position of the paddle tip above the undisturbed surface. (b) Geometry of paddle in the frame of the moving water. The wetted length l_w shown here is the length of the paddle beneath the undisturbed free surface.

is small we can approximate the equation as

$$(M + m)\ddot{z} = -mg + F - \mu\dot{z}, \quad (2)$$

in which we define the effective mass M by $\mathcal{I} = (M + m)a^2$.

The Reynolds number based on the depth of the water is large, and is still large even if the length scale is taken as the depth of the paddle (~ 5 mm),

$$Re = \frac{UH}{\nu} \sim 10^4 - 10^5. \quad (3)$$

The relevance of surface tension is seen from the Weber number

$$We = \frac{\rho U^2 z}{\sigma} \sim 100, \quad (4)$$

in which we take z to be a typical depth of the paddle. Since this is large, inertia is much more important than surface tension forces which we will henceforth discount.

The remaining forces from the water might be separately described as buoyancy (hydrostatic pressure), hydrodynamic pressure and hydraulic drag. The buoyancy is simply the weight of the water displaced by the wedge shape of the paddle $\rho g W z^2 / 2 \tan \alpha$, which we will see is negligible in comparison with the hydrodynamic force $\rho U^2 W z$ at the speeds ($U \sim 1 \text{ m s}^{-1}$) of interest. The Froude number U^2 / zg is large, and this is indicative of the fact that buoyancy is not the principal force from the water; for simplicity, we will therefore neglect any effects of gravity on the water.

The vertical force of the water on the paddle can therefore be written as

$$F = F_n \cos \alpha - F_t \sin \alpha \quad (5)$$

where F_n is the force acting in the normal direction on the paddle and F_t is the force acting in the downwards tangential direction. We expect the normal force here to be due to the hydrodynamic pressure and the tangential force to be due to hydraulic drag, and for the purposes of the current experiment we expect the normal force producing *hydrodynamic lift* to be the most important.

We begin by following the approach used with some success to model skipping stones [2, 3, 13], in which the force depends on the relative velocity of the stone or paddle, and its depth in the water. At high Reynolds number, such a force takes the form

$$F = \frac{1}{2}C\rho|\mathbf{U}|^2A, \quad (6)$$

in which $\mathbf{U} = -U\mathbf{i} + \dot{z}\mathbf{k}$ is the relative velocity, A is the cross sectional area of the object, and C is a coefficient.

For the context of an object entering the water we expect that the relevant area should be the cross sectional area in the direction of motion. For an object traveling at an angle $\beta = -\tan^{-1}(\dot{z}/U)$ to the horizontal (relative to stationary water) we can therefore expect (figure 8),

$$F_n = \frac{1}{2}\rho|\mathbf{U}|^2A = \frac{1}{2}\rho|\mathbf{U}|^2Wl_w \sin(\alpha + \beta), \quad (7)$$

where $l_w = -z/\sin \alpha$ is the ‘wetted length’ of the paddle, and W is the width of the paddle. Note that the wetted length here is taken as being the length of the paddle beneath the undisturbed water surface: this definition will need to be refined later. This expression for the force was deduced from experiments conducted by Rosellini et al. [13], but can be understood in terms of the geometry of figure 8. The coefficient C in (6) is therefore 1 in this case, provided we choose the correct definition of the area A .

(7) can be written as

$$F_n = -\frac{1}{2}\rho W(U^2 + \dot{z}^2)^{1/2}z(U - \cot \alpha \dot{z}), \quad (8)$$

and when \dot{z} is much smaller than U (an approximation which is appropriate for our experiment) this force becomes linear in z and we have, ignoring friction for the moment,

$$(M + m)\ddot{z} = -mg - \frac{1}{2}\rho W \cos \alpha U^2 z, \quad (9)$$

This is a linear spring equation which oscillates about the steady position

$$z_0 = -\frac{2mg}{\rho W \cos \alpha U^2}, \quad (10)$$

with frequency

$$\omega = \left[\frac{\rho W \cos \alpha U^2}{2(M + m)} \right]^{1/2} \text{ rad s}^{-1}. \quad (11)$$

However this assumes that the paddle is always in the water, whereas of course if it goes above the surface it no longer feels the effect of the water at all. So we can write the equation of motion as

$$\ddot{z} = \omega^2 z_0 - \omega^2 z \mathcal{H}(-z), \quad (12)$$

where $\mathcal{H}(-z)$ is the Heaviside function (0 for $z > 0$ and 1 for $z < 0$). If the oscillations have large enough amplitude they will go above the surface and have projectile motion in the air,

$$\ddot{z} = \omega^2 z_0. \quad (13)$$

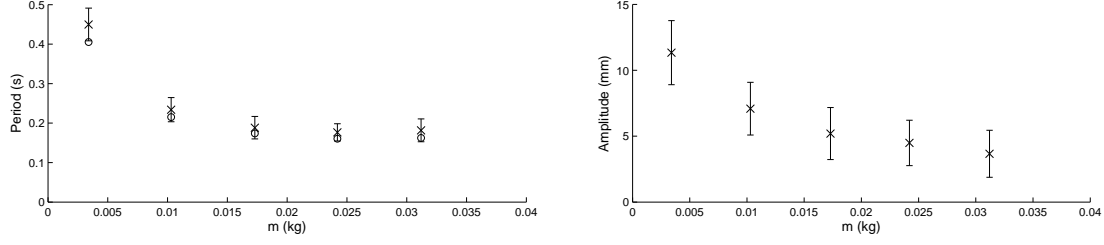


Figure 9: Average period and amplitude of oscillations over a 45 s time interval, for water speed $U = 1.5 \text{ m s}^{-1}$, with 5 different masses. The error bars show one standard deviation either side of the mean (crosses), and the circles show the theoretical prediction for the period (17) if the amplitude is fitted to the observed average for each mass.

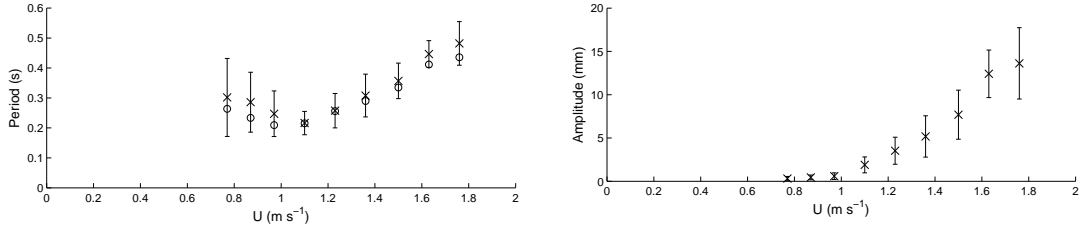


Figure 10: Average period and amplitude of oscillations over a 45 s time interval, for paddle mass $m = 3.4 \text{ g}$, with different water speeds. The error bars show one standard deviation either side of the mean (crosses), and the circles show the theoretical prediction for the period (17) if the amplitude is fitted to the observed average for each speed.

The velocity is zero at the maximum height H , and the path is parabolic,

$$z = H + \frac{1}{2}\omega^2 z_0 t^2, \quad (14)$$

entering the water at time $t_c = (-2H/\omega^2 z_0)^{1/2}$ with velocity $v = -(-2H\omega^2 z_0)^{1/2}$. After this the motion is sinusoidal

$$z = z_0 - z_0 \cos \omega(t - t_c) - (-2H z_0)^{1/2} \sin \omega(t - t_c), \quad (15)$$

and will exit the water again at a time

$$t = t_c + \frac{2}{\omega} \left[\pi - \tan^{-1} \left(-\frac{2H}{z_0} \right)^{1/2} \right]. \quad (16)$$

The period of the whole oscillation (for $H > 0$) is therefore

$$\frac{2}{\omega} \left[\pi - \tan^{-1} \left(-\frac{2H}{z_0} \right)^{1/2} + \left(-\frac{2H}{z_0} \right)^{1/2} \right]. \quad (17)$$

Unlike a simple harmonic oscillator the frequency now depends on the amplitude of the motion, which is seen from the solutions to be

$$H - z_0 + (z_0^2 - 2Hz_0)^{1/2}. \quad (18)$$

Thus, if we know the amplitude of oscillations we should be able to predict the period or frequency. By taking the observed amplitudes of the oscillations in the experiment we can estimate what the periods should be, and the results of such calculations are shown in figures 9 and 10.

However, there is nothing in (12) which sets the amplitude of oscillations - they are purely determined by the initial conditions. Moreover, if we introduce the frictional term $-\mu\dot{z}$ any initial conditions will eventually decay towards the steady state. The same will be true for any force $F(z)$, since the change in energy over a collision starting with $z = 0$ at $t = t_{IN}$ and ending at $z = 0$ at $t = t_{OUT}$ is

$$\left[\frac{1}{2}(M + m)\dot{z}^2 \right]_{t_{IN}}^{t_{OUT}} = \int_{t_{IN}}^{t_{OUT}} -mg \dot{z} dt + \int_{t_{IN}}^{t_{OUT}} F_n(z) \cos \alpha \dot{z} dt - \int_{t_{IN}}^{t_{OUT}} \mu \dot{z}^2 dz \quad (19)$$

$$= - \int_{t_{IN}}^{t_{OUT}} \mu \dot{z}^2 dz < 0. \quad (20)$$

Thus any force which is a function only of z will result in damped oscillations. In the experiments we observe that there are oscillations which continue indefinitely, and their amplitude naturally selects itself. Thus in the equation of motion (2) there must be some destabilizing process which drives oscillations and it is then the balance between this and the stabilizing effects of friction which determines their amplitude.

We see that approximating the force in (8) by the linear force in (9) cannot explain the observed oscillations. By including the full force (8) we also find that the additional nonlinear terms have a dissipative effect; figure 11 shows the solution to the equation of motion with F taken as the vertical component of (8):

$$(M + m)\ddot{z} = -mg - \frac{1}{2}\rho W \cos \alpha (U^2 + \dot{z}^2)^{1/2} z (U - \cot \alpha \dot{z}). \quad (21)$$

For comparison the solution is compared with the solution when the horizontal velocity $U = \dot{x}$ is also allowed to vary due to the horizontal component of (8),

$$(M + m)\ddot{x} = -\frac{1}{2}\rho W \sin \alpha (U^2 + \dot{z}^2)^{1/2} z (U - \cot \alpha \dot{z}), \quad (22)$$

as is appropriate for the usual case of a skipping stone.

It seems that the force in (8) is inadequate to describe our observations because it does not contain any destabilizing effect. This is more evident in the experiments if we look at the vertical velocity of the paddle while it is continually skipping. Figure 12 shows this velocity alongside the displacement of the paddle, and it is seen that the vertical velocity is on average larger when it leaves the water than when it enters the water. Plotting the exit velocity against entry velocity for each collision (figure 13) we find that the collisions are super-elastic, meaning the coefficient of normal restitution $e = V_{OUT}/V_{IN}$ is larger than 1, and this effect grows as the speed of the water is increased.

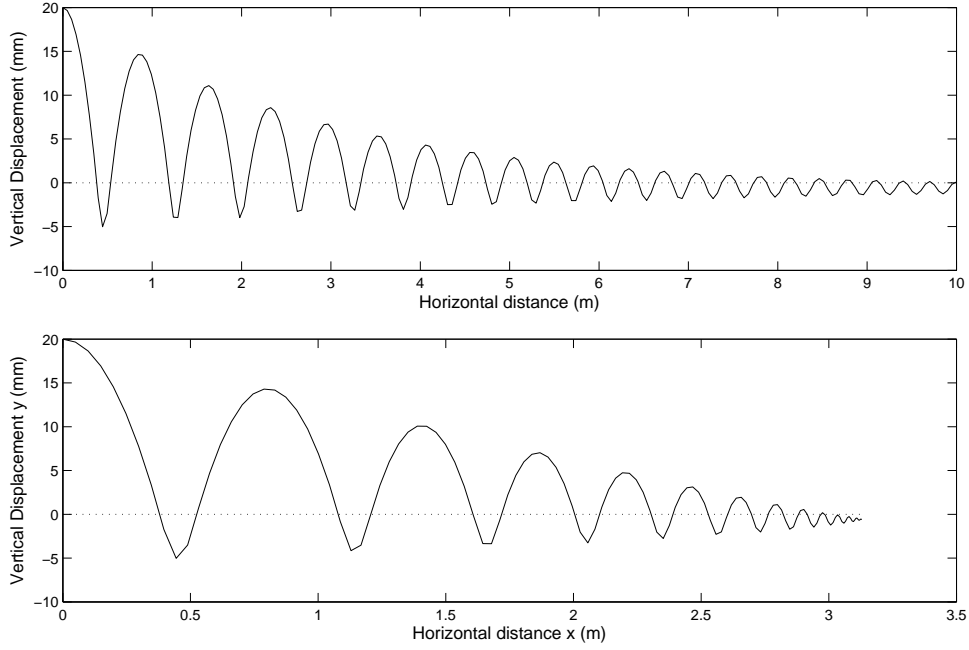


Figure 11: (a) Trajectory of a light stone skipping using Rosellini et al.'s [13] parameterisation of the hydrodynamic force on the stone, with the horizontal velocity imposed as $U = 1.5 \text{ m s}^{-1}$ (21). In this case the distance x is equivalent to time t . The dashed line shows the undisturbed water level. The vertical force on the particle varies through the collision, causing the bounces to decay. (b) The same but with the horizontal velocity unconstrained, so that the position $x(t)$ is determined from (22). The horizontal velocity starts at 1.5 m s^{-1} and decreases to zero due to the force of the water.

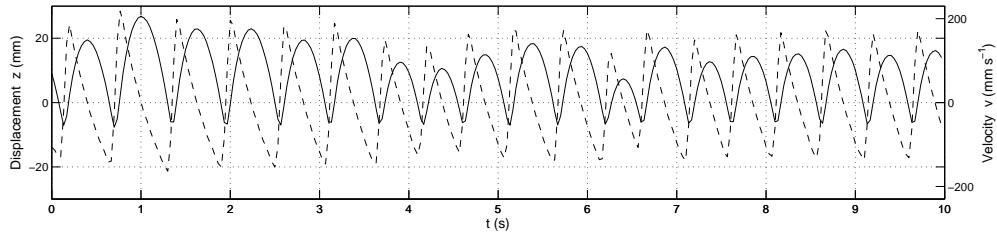


Figure 12: Vertical displacement (solid, left axis) and vertical velocity (dashed, right axis), for the 10 cm wide paddle with mass 3.4 g and water speed $U = 1.76 \text{ m s}^{-1}$.

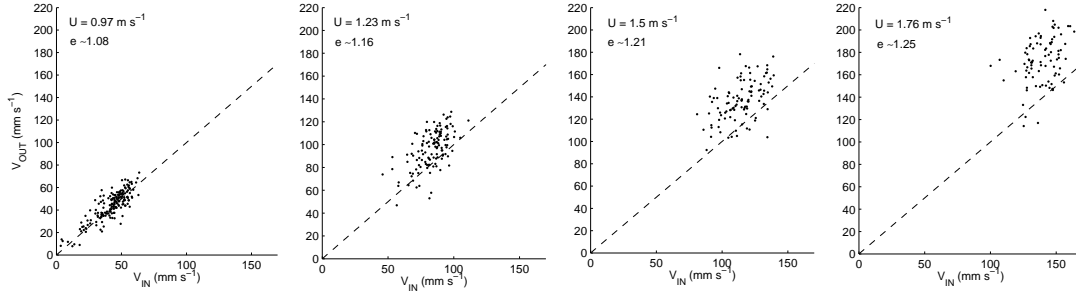


Figure 13: Maximum downward velocity (V_{IN}) plotted against the immediately following maximum upward velocity (V_{OUT}) for each oscillation in a 45 second time interval, for the 10 cm wide paddle with mass 3.4 g and different water speeds U as shown. V_{IN} and V_{OUT} presumably occur somewhere close to where the paddle enters the water when it is bouncing. The values given for the effective coefficient of restitution e are the average of V_{OUT}/V_{IN} .

So the paddle gains kinetic energy during collisions, and this kinetic energy gain must be balanced by the loss due to friction during the air-borne motion. In the next section we discuss what aspects of the hydrodynamic force that can cause this are missing from (8).

3.1 The rise of water ahead of the paddle

The discussion of the previous section leads us to believe that the force from the water cannot be simply a function of the depth beneath the undisturbed surface level, but varies in a more complicated way throughout the course of the collision. The reason for this is due to the water ahead of the paddle piling up there, so that the effective depth, and therefore the wetted length l_w are not just dependent on z (figure 14). This fact was realised by Wagner [20] who noted that the water around an impacting object rises up to meet it, so that the area of the water wetting the object is like an expanding plate. Thus the region over which the water pressure is appreciable is not simply the region beneath the undisturbed surface, but a larger region. It is necessary to determine what this region is and how it varies during the course of the collision in order to calculate the force on the object.

A similar mechanism has been used to explain anomalous (larger than 1) values of the coefficient of normal restitution in oblique elastic collisions - the surface deforms asymmetrically so that the effective normal for the collision is tilted [11, 10].

We can still think of the force as in (7), but we need to take more care in determining l_w . There have been some experimental studies of the water rise ahead of a planing hull [12] which attempt to determine the wetted length as a function of the hull's width and trim angle (α). However we expect the *variation* of this wetted length to be important and in particular we need to consider how l_w varies during the course of the collision. It is not entirely clear how to do this in a simple way, so we begin by considering a two-dimensional shallow water model.



Figure 14: Snapshots of paddle and free surface over the course of a collision showing the evolution of the splash-up region ahead of the paddle.

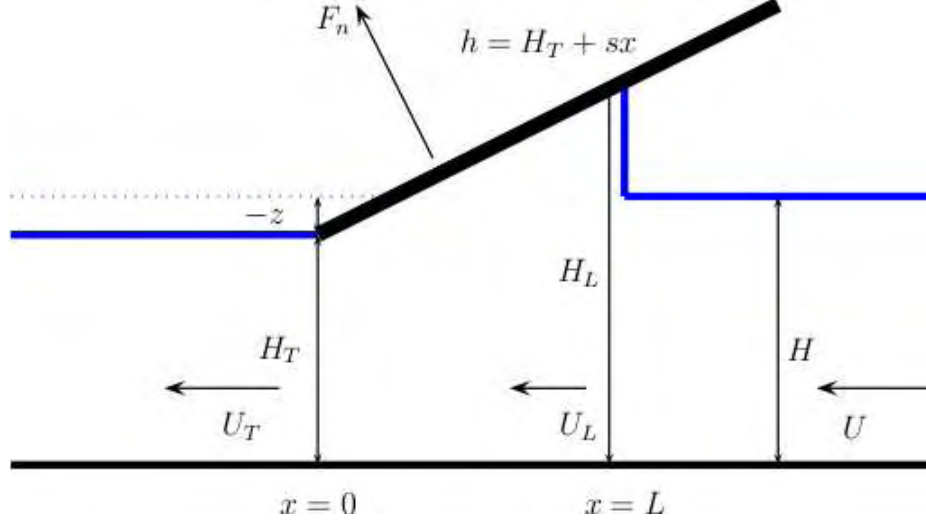


Figure 15: Shallow water model setup with the definition of the water depths H , H_L and H_T , and the position of the shock $x = L(t)$. The upstream conditions H , U are prescribed constants, but the other variables are changing over the course of the collision.

4 Shallow water model

The situation we consider is shown in figure 15; it is similar to the steady case modeled by Tuck & Dixon [17, 14] for the shallow flow beneath a surfboard. The difference between their model and ours is that we allow the height of the paddle to change and we treat the jet which is ejected ahead of the paddle as producing a leading ‘wedge’ of water. We treat this splash-up region ahead of the paddle as a shock, similar to a tidal bore on a river estuary. The shallow water equations that apply beneath the paddle are, neglecting gravity,

$$h_t - (uh)_x = 0, \quad (23)$$

$$\bar{\rho}(u_t - uu_x) = p_x, \quad (24)$$

in which h is the water depth, u is the velocity in the negative x direction, p is the pressure and $\bar{\rho} = \rho W$ is the two-dimensional density.

The upstream values H and U are steady, and the values at the trailing edge are labeled $h = H_T$ and $u = U_T$ (these will vary in time). At the shock $x = L(t)$ we label $h = H_L$ and $u = U_L$; the differential equations (23) and (24) don’t apply here since there is a discontinuity, but mass and momentum must still be conserved so we require jump

conditions,

$$-(\bar{\rho}H_L - \bar{\rho}H)\dot{L} = (\bar{\rho}H_L U_L - \bar{\rho}H U), \quad (25)$$

$$-(\bar{\rho}H_L U_L - \bar{\rho}H U)\dot{L} = (\bar{\rho}H_L U_L^2 + p_L H_L - \bar{\rho}H U^2). \quad (26)$$

Note that the pressure on the upstream face of the shock is zero (atmospheric), since we are neglecting gravity.

If we know the position of the trailing edge of the paddle z above the undisturbed surface then $H_T = H + z$ and the depth of water beneath the paddle (for $0 < x < L(t)$) is

$$h(x, t) = H_T + sx, \quad (27)$$

where $s = \tan \alpha$ is the slope of the paddle.

The velocity $u(x, t)$ under the paddle is therefore related to the velocity at the trailing edge by mass conservation (23),

$$u(x, t) = \frac{u_T H_T + \dot{z}x}{H_T + sx} = U + \frac{(H + z)\Delta U - (sU - \dot{z})x}{H + sx + z}, \quad (28)$$

in which we define $\Delta U = U_T - U$ as the change in water speed from upstream to downstream of the paddle. The pressure is obtained from integrating the momentum equation (24), also in terms of ΔU ;

$$\begin{aligned} \frac{p(x, t)}{\bar{\rho}} = & \frac{(sU + s\Delta U - \dot{z})^2 x}{s(H + sx + z)} - \frac{(sU + s\Delta U - \dot{z})^2 x^2}{2(H + sx + z)^2} + \frac{\ddot{z}x}{s} \\ & + \frac{(H + z)(s\Delta \dot{U} - \ddot{z}) + \dot{z}(sU + s\Delta U - \dot{z})}{s^2} \log \left(1 + \frac{sx}{H + z} \right), \end{aligned} \quad (29)$$

the dots all denoting time derivatives.

Combining the jump conditions (25) and (26) gives

$$p_L(H_L - H) = \bar{\rho}H(U - U_L)^2, \quad (30)$$

in which we know

$$H_L = H + z + sL. \quad (31)$$

(25) and (30), with H_L , u_L and p_L coming from (31), (28) and (29), give two equations for the unknowns L and ΔU which also involve z and its time derivatives, and must therefore be coupled with the equation of motion for the paddle

$$(M + m)\ddot{z} = -mg + F_n \cos \alpha - \mu\dot{z}, \quad (32)$$

with F_n now given by

$$\begin{aligned} \frac{F_n(t)}{\bar{\rho}} = & \int_0^{L(t)} \frac{p(x, t)}{\bar{\rho}} dx = \frac{(sU + s\Delta U - \dot{z})^2 L^2}{2s(H + sL + z)} + \frac{\ddot{z}L^2}{2s} \\ & + ((H + z)(s\Delta \dot{U} - \ddot{z}) + \dot{z}(sU + s\Delta U - \dot{z})) \left\{ \frac{H + sL + z}{s^3} \log \left(1 + \frac{sL}{H + z} \right) - \frac{L}{s^2} \right\}, \end{aligned} \quad (33)$$

The combination (25), (30), (31), (28), (29), (32) and (33) provides 3 ordinary differential equations for the unknowns z , L and ΔU . The initial conditions for the start of a collision are

$$z = 0, \quad \dot{z} = V_{IN}, \quad L = 0, \quad \Delta U = 0, \quad \text{at } t = 0. \quad (34)$$

(29) provides an expression for p_L , and from (30) we also have

$$\frac{p_L}{\bar{\rho}} = \frac{H((H+z)\Delta U - (sU - \dot{z})L)^2}{(H + sL + z)^2(sL + z)}. \quad (35)$$

Combining these we get an expression for $\Delta \dot{U}$,

$$\begin{aligned} \Delta \dot{U} \frac{H+z}{s} \log \left(1 + \frac{sL}{H+z} \right) \\ = \frac{H((H+z)\Delta U - (sU - \dot{z})L)^2}{(H + sL + z)^2(sL + z)} - \frac{(sU + s\Delta U - \dot{z})^2 L}{s(H + sL + z)} + \frac{(sU + s\Delta U - \dot{z})^2 L^2}{2(H + sL + z)^2} \\ - \frac{\dot{z}(sU + s\Delta U - \dot{z})}{s^2} \log \left(1 + \frac{sL}{H+z} \right) - \frac{\ddot{z}}{s^2} \left(sL - (H+z) \log \left(1 + \frac{sL}{H+z} \right) \right), \end{aligned} \quad (36)$$

which can be substituted into the force (33) in the equation of motion to find \ddot{z} . After rearrangement we obtain

$$\begin{aligned} \ddot{z} \left[\frac{M+m}{\rho W \cos \alpha} - \frac{L^2}{2s} + \frac{(H + sL + z)L \log \left(1 + \frac{sL}{H+z} \right) - sL^2}{s^2 \log \left(1 + \frac{sL}{H+z} \right)} \right] \\ = -\frac{mg}{\rho W \cos \alpha} - \frac{\mu}{\rho W \cos \alpha} \dot{z} + \frac{(sU + s\Delta U - \dot{z})^2 L^2}{2s(H + sL + z)} + \frac{(H + sL + z) \log \left(1 + \frac{sL}{H+z} \right) - sL}{s \log \left(1 + \frac{sL}{H+z} \right)} \\ \times \left[\frac{H((H+z)\Delta U - (sU - \dot{z})L)^2}{(H + sL + z)^2(sL + z)} - \frac{(sU + s\Delta U - \dot{z})^2 L}{s(H + sL + z)} + \frac{(sU + s\Delta U - \dot{z})^2 L^2}{2(H + sL + z)^2} \right]. \end{aligned} \quad (37)$$

We are left with the substantially shorter equation for \dot{L} ,

$$(sL + z)\dot{L} = -H\Delta U - (U + \Delta U)z - \dot{z}L. \quad (38)$$

These equations apply only when the paddle is in the water; that is, $L > 0$. When $L < 0$, the paddle follows a ballistic trajectory given by

$$(M + m)\ddot{z} = -mg - \mu\dot{z}. \quad (39)$$

The equations (36), (37) and (38) must be solved numerically, and a typical solution is shown in figure 16. Potential singularities in the equations are avoided by switching to equation (39) whenever L is smaller than some small threshold.

It is found that the long term behaviour is independent of the initial condition - at very low water speeds there is a steady state which is stable, but at higher water speeds there is a natural amplitude oscillatory mode, towards which any initial condition will evolve.

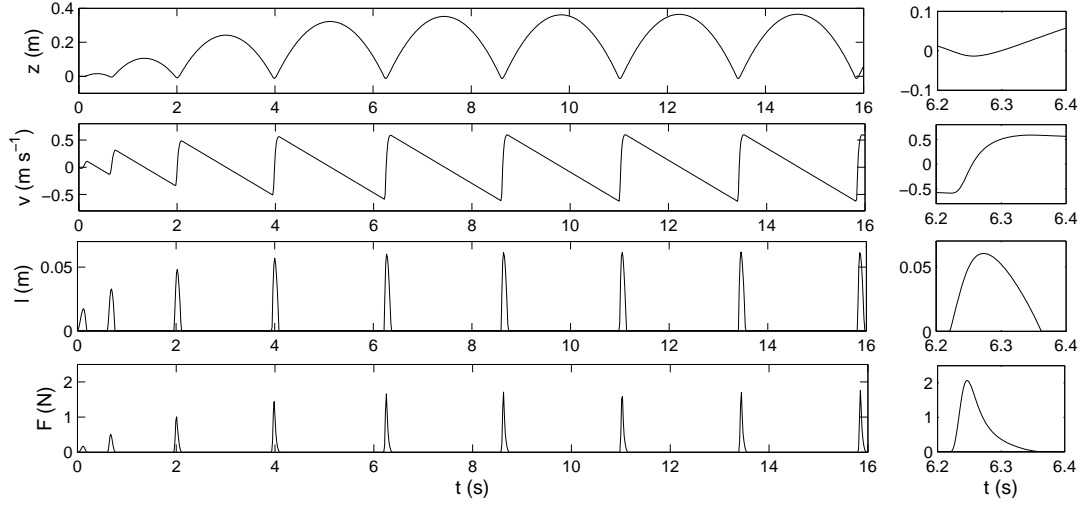


Figure 16: Solution to the shallow water model, (36), (37) and (38) with no friction or mass loss terms, showing the vertical position z , velocity $v = \dot{z}$, position of shock L and vertical force F . The depth is 6 cm, the water speed is 1.5 m s^{-1} , and the paddle mass is 3.4 g. The initial condition is to start from rest on the undisturbed surface $z = 0$. On the right is an enlarged view of one of the collisions.

4.1 Flow around the sides of the paddle

It is evident from figure 16 that these oscillations are considerably larger than those observed in the experiment, with a consistently longer period. One reason for this is undoubtedly the fact that the current model is two dimensional and the water is forced to flow under the paddle, whereas what actually happens is that a substantial fraction of the water passes around the paddle to the sides.

One way of including this mass loss into the two dimensional model is to treat it as occurring at the shock front, so that instead of conserving mass the shock actually allows a certain amount to be lost. If the position of the shock is $x = L$, then the area of the triangular wedge which has built up ahead of the paddle is $sL^2/2$. We might expect on dimensional grounds that the rate at which mass is lost to the sides is proportional to this area and the speed of the water, so that the mass loss is

$$-\lambda \rho s L^2 U. \quad (40)$$

Then the modified jump condition (25) becomes

$$-\rho W(H_L - H)\dot{L} = \rho W(H_L U_L - H U) + \rho s L^2 U. \quad (41)$$

With this modification the expression for the pressure (35) acquires an extra term

$$\frac{p_L}{\bar{\rho}} = \frac{H((H + z)\Delta U - (sU - \dot{z})L)^2}{(H + sL + z)^2(sL + z)} + \lambda \frac{sL^2 U(H\Delta U + (U + \Delta U)z + \dot{z}L)}{W(H + sL + z)(sL + z)}, \quad (42)$$

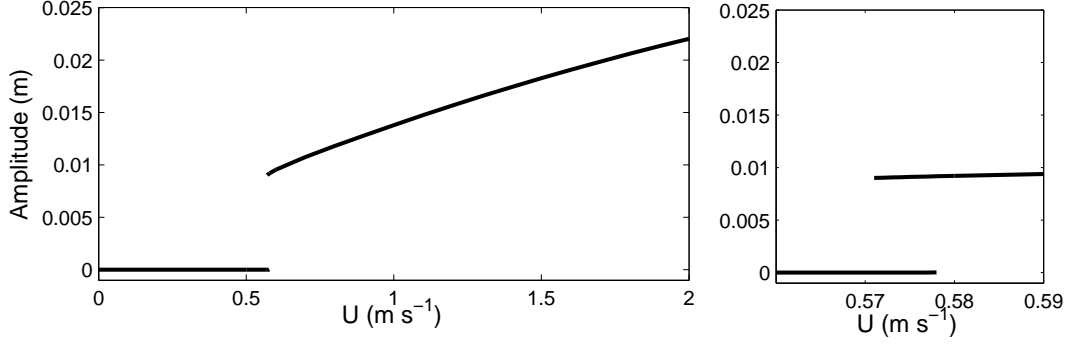


Figure 17: Bifurcation diagram for shallow water model showing oscillation amplitudes as a function of water speed. On the right is an enlarged view of the transition. The steady state becomes linearly unstable at around 0.578 and the upper (bouncing) branch loses stability at around 0.571. $m = 3.4$ g, $\mu = 0.2$ kg s⁻¹ and $\lambda = 1$.

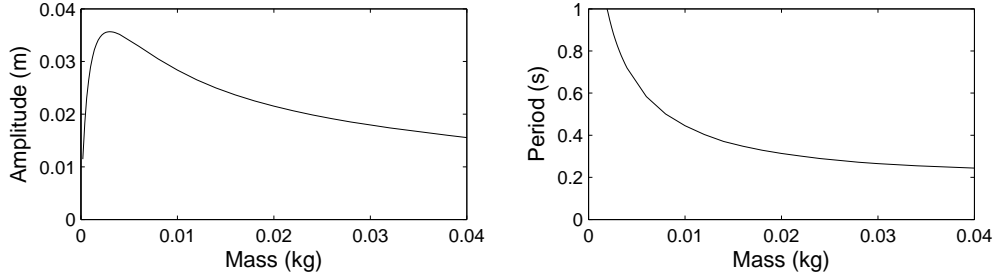


Figure 18: Amplitude and period of oscillations for the shallow water model with constant water speed $U = 1.5$ m s⁻¹, and varying mass m . $\mu = 0.2$ kg s⁻¹ and $\lambda = 1$.

which also therefore appears in (36) and (37). (38) becomes

$$(sL + z)\dot{L} = -H\Delta U - (U + \Delta U)z - \dot{z}L - \lambda \frac{sL^2U}{W}. \quad (43)$$

Including this effect reduces the size of the leading wedge of water and results in a smaller impulse from the water, so that the height of bounces is not as large. Whereas (36), (37) and (38) without the mass loss terms only have a steady state in which z is positive (the trailing edge of the paddle is *above* the undisturbed surface), including mass loss allows it to have a steady planing state which is (perhaps more reasonably) beneath the undisturbed surface. This steady state is stable at low water speeds, but becomes unstable above a threshold value, and any initial oscillations then grow until the paddle leaves the water surface. Again, the eventual oscillatory mode is independent of the initial conditions. This allows us to examine how the amplitude and period of oscillations behave as the water speed or mass vary; figure 17 shows the amplitude as the water speed is increased, and figure 18 shows how the amplitude and period vary with mass for a given speed. These should be compared qualitatively with figures 9 and 10.

The steady branch of the bifurcation diagram in figure 17 becomes linearly unstable at

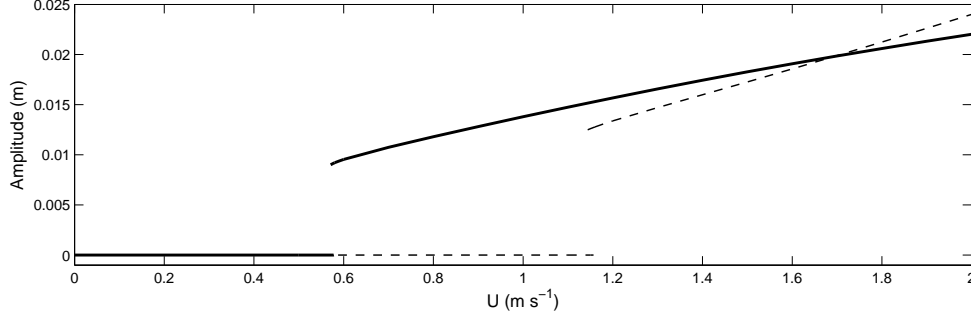


Figure 19: Bifurcation diagram for shallow water model showing oscillation amplitudes as a function of water speed for two different masses; 3.4 g (solid) and 20 g (dashed). The threshold speed for oscillations is higher for the larger mass. $\mu = 0.2 \text{ kg s}^{-1}$ and $\lambda = 1$.

a slightly larger water speed than where the oscillatory branch ceases to exist; this suggests the bifurcation is subcritical and there may be an unstable branch between the two. The oscillations of the upper branch are always leaving the water surface - there appear to be no steady oscillations which remain in the water. Above the threshold the amplitude increases with water speed; this is what the experiments show, except that there is no detectable jump in amplitude at which the oscillations start. The model suggests however that just below the threshold any initial oscillations will decay very slowly, so that with a small amount of noisy forcing (such as the effect of the wake in the tank), any oscillations below the threshold would be hard to get rid of. It is possible that the unavoidable roughness in the water surface might act to smooth out a sharp bifurcation in the measurements.

Figure 19 compares the bifurcation diagram for two different masses; for a larger mass the threshold velocity at which the steady state loses stability is larger, as the experiments (figure 4) suggest ought to be the case. However, it also seems that if the water speed is large enough the heavier paddle can oscillate more, thus reversing the trend seen in the experiments. It is not clear whether this is an artifact of the model or if the experiments did not reach large enough speeds to detect it.

When the paddle is really skipping and spending the majority of its time in the air, the time in the water can be viewed as a bounce, just like the bounce of an elastic ball. One way of thinking about the amplitude of the skips in this case is to consider the change in velocity (or kinetic energy) effected by each discrete bounce. Just like an elastic bounce with coefficient of normal restitution e will return an incident velocity V_{IN} at velocity $V_{OUT} = eV_{IN}$, each collision with the water will return an incident velocity V_{IN} at a different velocity which we can calculate from our model. The equilibrium amplitude which is reached is when this increase in velocity or energy is matched by the decrease due to friction during the airborne part of the motion, which we can similarly treat as providing a particular incident velocity V_{IN} (for the next collision) for any given V_{OUT} . If we plot V_{OUT} as a function of V_{IN} for the bounce and V_{IN} as a function of V_{OUT} for the ballistic motion, the steady state will be where these two curves intersect (if they do not intersect there is no steady oscillation). Provided there is one intersection, for any initial velocity we can trace a ‘staircase’ trajectory between the curves so that any initial condition will grow

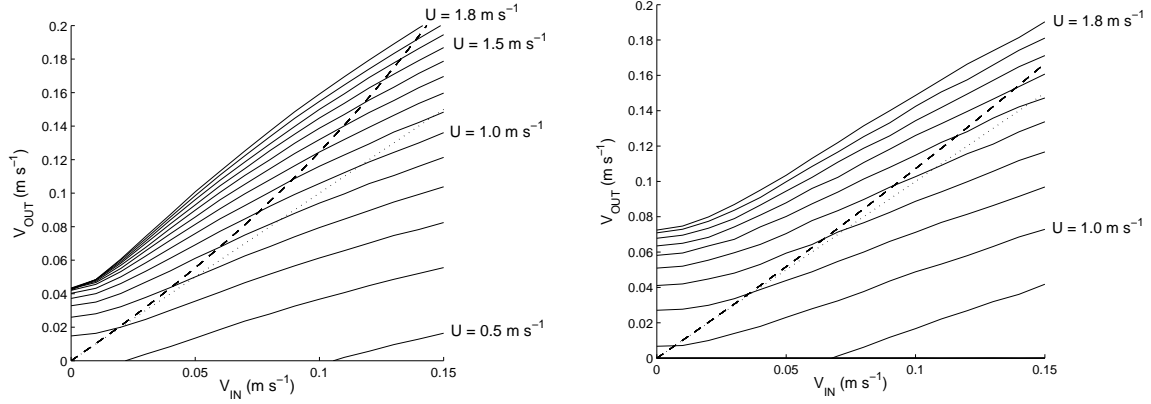


Figure 20: Exit velocity V_{OUT} as a function of entry velocity V_{IN} (solid lines) for collisions in the shallow water model, with $\mu = 0.1 \text{ kg s}^{-1}$ and $\lambda = 1$, $\alpha = 20^\circ$, and (a) $m = 3.4 \text{ g}$, (b) $m = 10 \text{ g}$; The different curves correspond to different water speeds U at intervals of 0.1 m s^{-1} . The dashed line shows the entry velocity as a function of exit velocity for ballistic motion with friction μ , and the dotted line is $V_{OUT} = V_{IN}$.

or decay towards the equilibrium.

Two sets of such curves are shown in figure 20 for two different masses. The collisional curves are different for different water speeds whereas the airborne curves are naturally independent of the water. For small enough U there is no intersection and therefore no steady bouncing state. As U is increased the point of intersection has larger V_{OUT} , corresponding to a larger amplitude. Unfortunately the shape of the curves depends on both U and m , and each curve has to be computed separately; there is no simple scaling which determines the intersection for different U or m .

5 A simplified shallow water model

Whilst the shallow water model presented above does a reasonable job of recreating some of the experimental trends it is not quite obvious from (36), (37) and (38) what is going on in the collision, nor how to generalise the model to deep water. If the displacement of the water surface is small, then we can attempt to extract from the equations a simpler model which includes fewer of the terms. We expect z , L and ΔU and their time derivatives all to be small in this case.

Assuming that all these terms are equally small, we find the jump conditions both become steady at leading order, and can only be satisfied if $z = 0$. We therefore expect z to be smaller than L , and the mass jump condition suggests $L^2 \sim z$. We thus look for the leading order terms in the equations when L is small and z and ΔU are quadratically small (that is, in comparison to H and U respectively; formally we treat $L/H = O(\varepsilon)$, $z/H = O(\varepsilon^2)$ and $\Delta U/U = O(\varepsilon^2)$). The equations (36), (43) and (42) become

$$-2U\Delta U - \frac{s^2 L^2 U^2}{2H^2} - \frac{U^2 z}{H} = O(\varepsilon^3), \quad (44)$$

$$sL\dot{L} = -H\Delta U - Uz - \lambda \frac{sL^2U}{W} + O(\varepsilon^3), \quad (45)$$

$$(M + m)\ddot{z} = -mg - \mu\dot{z} + \frac{sL^2U^2}{2H} + O(\varepsilon^3), \quad (46)$$

the first two of which simplify to

$$sL\dot{L} = -\frac{Uz}{2} + \frac{s^2L^2U}{4H} - \lambda \frac{sL^2U}{W} + O(\varepsilon^3). \quad (47)$$

It is worth considering what this approximation means in terms of the physical conservation laws on which the model is based. The leading order momentum jump condition is steady, so force balance occurs across the shock;

$$\frac{p_L}{\bar{\rho}} H_L = HU^2 - H_L U_L^2. \quad (48)$$

Moreover the u_t term in the momentum equation is smaller than the uu_x term, so that the first order pressure is determined by the steady Bernoulli equation

$$\frac{p_L}{\bar{\rho}} = \frac{1}{2}U_T^2 - \frac{1}{2}U_L^2. \quad (49)$$

The h_t term in the mass conservation equation beneath the paddle is also small, so that at leading order the mass flux beneath the paddle is uniform, $H_L U_L = H_T U_T$. The only balance which is not steady is therefore the mass balance across the shock at $x = L$ (43),

$$(H_L - H)\dot{L} = HU - H_L U_L - \lambda \frac{sL^2U}{W}. \quad (50)$$

The mass flux $H_L U_L$ comes from combining the two momentum balances to give, up to quadratically small terms,

$$H_L U_L = U \left(\frac{2HH_L H_T^2}{H_T^2 + H_L^2} \right)^{1/2}, \quad (51)$$

but perhaps a better way of writing (50) is (correct to quadratic terms),

$$sL\dot{L} = -zU - H\Delta U - \lambda \frac{sL^2U}{W}, \quad (52)$$

in which ΔU follows from the *global* force balance

$$\frac{sF_n}{\bar{\rho}} = HU^2 - H_T U_T^2, \quad (53)$$

$$= -zU^2 - 2HU\Delta U. \quad (54)$$

The leading order force follows from integrating Bernoulli's equation for the pressure,

$$\frac{F_n}{\bar{\rho}} = \frac{sL^2U^2}{2H}, \quad (55)$$

so that combining (52) and (53), and the equation of motion, we recover (46) and (47).

The equation (52) (or (47)) for L can be thought of as mass conservation for the triangular wedge of water ahead of the paddle - the three terms on the right can be viewed respectively as the rate of mass flow into the cross sectional area of the paddle facing the fluid, the effect of the force of the paddle acting to slow down the whole water column, and the mass lost to the sides of the paddle.

5.1 Linear Stability

The simplified model (43) and (47) is in fact linear if we use the area of the water wedge $A = sL^2/2$ as a variable,

$$(M + m)\ddot{z} = -mg + \rho W \cos \alpha \frac{AU^2}{H} - \mu \dot{z}, \quad (56)$$

$$\dot{A} = -\frac{zU}{2} + \frac{sAU}{2H} - \lambda \frac{AU}{W}. \quad (57)$$

(56) is the same as our original linear equation of motion (9), except that the force now depends on the area of the leading wedge A/H rather than directly on the depth z . So we have essentially the same model, but now with a separate evolution equation for the ‘wetted length’ A/H .

There is a steady (planing) state

$$A_0 = \frac{mgH}{\rho W \cos \alpha U^2}, \quad z_0 = -\left(\frac{2\lambda}{W} - \frac{s}{H}\right) \frac{mgH}{\rho W \cos \alpha U^2}, \quad (58)$$

and we can examine its stability by writing the equations in the matrix form

$$\begin{pmatrix} \dot{A} \\ \dot{z} \\ \ddot{z} \end{pmatrix} = \begin{pmatrix} b & -a & 0 \\ 0 & 0 & 1 \\ c & 0 & -d \end{pmatrix} \begin{pmatrix} A - A_0 \\ z - z_0 \\ \dot{z} \end{pmatrix}, \quad (59)$$

where

$$a = \frac{U}{2}, \quad b = \frac{sU}{2H} - \frac{\lambda U}{W}, \quad c = \frac{\rho W \cos \alpha U^2}{M + m} \frac{1}{H}, \quad d = \frac{\mu}{M + m}. \quad (60)$$

Solutions of the form $e^{\lambda t}$ exist provided λ satisfies

$$\lambda^3 - (b - d)\lambda^2 - bd\lambda + ac = 0. \quad (61)$$

Since a and c are positive this clearly has one negative real root λ_0 and two other roots which could be real (in which case they are positive) or complex, but whose real part is of interest. If we write the complex roots as $\beta \pm i\alpha$ then

$$0 = (\lambda - \lambda_0)(\lambda - \beta - i\alpha)(\lambda - \beta + i\alpha) = \lambda^3 - (\lambda_0 + 2\beta)\lambda^2 + (\beta^2 + \alpha^2 + 2\beta\lambda_0)\lambda - (\beta^2 + \alpha^2)\lambda_0, \quad (62)$$

from which we require

$$\lambda_0 + 2\beta = b - d, \quad \beta^2 + \alpha^2 + 2\beta\lambda_0 = -bd, \quad (\beta^2 + \alpha^2)\lambda_0 = -ac. \quad (63)$$

If b is positive, the coefficient of λ^2 implies β must be positive (so the steady state is unstable), but if b is sufficiently negative we could have negative β (and therefore stability). The threshold for stability is (setting $\beta = 0$)

$$d > \frac{b}{2} + \left(\frac{b^2}{4} - \frac{ac}{b}\right)^{1/2}, \quad (64)$$

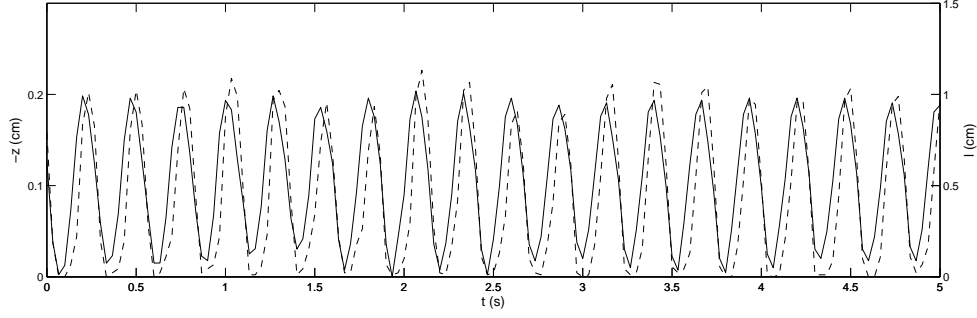


Figure 21: Time series of vertical position of paddle z (solid, left axis, larger values of $-z$ are deeper in the water), and wetted length of the paddle l (dashed, right axis). These are from measurements every $1/30$ of a second.

or in terms of the parameters

$$\frac{\mu}{M+m} > \frac{U}{2} \left(\frac{s}{2H} - \frac{\lambda}{W} \right) + \frac{U}{2} \left(\left(\frac{s}{2H} - \frac{\lambda}{W} \right)^2 - \frac{2\rho W \cos \alpha}{(M+m)H \left(\frac{s}{2H} - \frac{\lambda}{W} \right)} \right)^{1/2}. \quad (65)$$

Thus there is a critical level of friction needed in order to stabilise the system in the planing state, and the critical value increases linearly with the speed of the water U . If there is not sufficient friction the steady state will be unstable and perturbations will grow exponentially until they reach the surface (that is, once A reaches 0). The hydrodynamic force then switches off, and this non-linearity serves to set the amplitude of the bounces.

5.2 A shallow water experiment

As a qualitative test of the shallow water model, including the idealization of the splash up as a shock, we designed an alternative experiment which comprised water flow under gravity down an inclined plane. The flow was maintained by pumping water from a lower tank into an upper one which overflowed down the ramp into the lower one. The flow on the ramp was about 4 mm deep and travelling at about 60 cm s^{-1} . The same paddle was placed in this stream and, when light enough, was observed to oscillate up and down, coming clear of the water momentarily. The paddle was widened so as to take up the full width of the stream and minimise the flow around the sides. This made the oscillations greater, as the theory would suggest. When the paddle was heavy enough there was a steady state in which a large (several cm) wedge of fluid built up ahead, whereas when the paddle was light enough this wedge of fluid oscillated to and from the trailing edge.

The crucial factor in the model we have proposed is the fact that the wetted length (or equivalently in this case, the size of the leading wedge of water) does not respond instantaneously to the depth of the paddle in the water. The lag between changing the depth and increasing the wetted length is the essential mechanism in driving the oscillations.

By recording the position of the front of the leading wedge of water at the same time as the vertical position of the paddle we are able to verify this lag experimentally. Figure

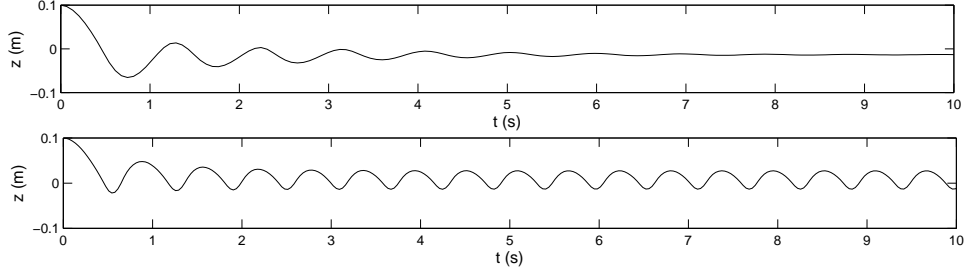


Figure 22: Paddle trajectories for the generalized deep water model (66) and (67), released from rest at 10 cm height, with water speed (a) 0.5 m s^{-1} (below threshold) and (b) 1 m s^{-1} (above threshold). $m = 10 \text{ g}$, $\mu = 0.2 \text{ kg s}^{-1}$ and $\lambda = 4$.

21 shows the position of the wetted front L and the depth of the paddle $-z$, and there is a clear lag between the two.

This experimental setup worked well and it would be good to conduct more tests with the same setup, particularly with a larger flow rate so that the speed of the water can be varied and its effect carefully monitored. A more precise bifurcation diagram might be found in this way than we were able to do in the rotating tank.

5.3 Generalization to deeper water

We now attempt to use the simplified model of the previous section to generalize to deeper water. (56) and (57) both contain the expression for the force $F_n = AU^2/H$ and this is the only place in which the water depth H appears. It seems natural to suppose that the water depth in this context represents the depth over which the presence of the paddle affects the water; for shallow water this is the whole depth, but for deeper water it might be expected to be on the scale of the horizontal perturbation (consider an analogy with surface gravity waves, when the phase speed changes from $(gH)^{1/2}$ to $(g/k)^{1/2}$ where k is the horizontal wave number). We therefore consider it appropriate to replace H in (56) and (57) by L ,

$$(M + m)\ddot{z} = -mg + \rho W \cos \alpha \frac{sLU^2}{2} - \mu\dot{z}, \quad (66)$$

$$sL\dot{L} = -\frac{zU}{2} + \frac{s^2UL}{4} - \lambda\frac{sL^2U}{2W}. \quad (67)$$

The steady state is now

$$L_0 = \frac{2mg}{\rho W \cos \alpha sU^2}, \quad z_0 = -\frac{mg}{\rho W \cos \alpha U} \left(\frac{2\lambda mg}{s\rho W^2 \cos \alpha U^2} - \frac{s}{2} \right). \quad (68)$$

Linearised perturbations about this again satisfy

$$\begin{pmatrix} \dot{L} \\ \dot{z} \\ \ddot{z} \end{pmatrix} = \begin{pmatrix} b & -a & 0 \\ 0 & 0 & 1 \\ c & 0 & -d \end{pmatrix} \begin{pmatrix} L - L_0 \\ z - z_0 \\ \dot{z} \end{pmatrix}, \quad (69)$$

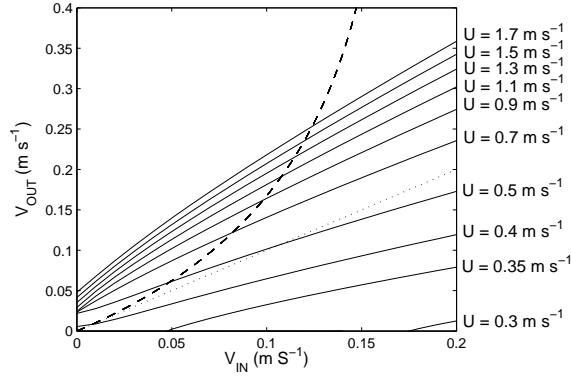


Figure 23: Exit velocity V_{OUT} as a function of entry velocity V_{IN} (solid) for deep water model (66) and (67). This is for a mass of 3.4 g and paddle of width 7 cm with mass loss $\lambda = 1$ and friction $\mu = 0.2 \text{ kg s}^{-1}$. The different curves correspond to different water speeds as labelled. The dashed line shows V_{IN} as a function of V_{OUT} as determined by resistance in the ballistic motion and the dotted line is the line $V_{IN} = V_{OUT}$.

but now with

$$a = \frac{\rho W \cos \alpha U^3}{4mg}, \quad b = \frac{\rho W \cos \alpha s^2 U^3}{8mg} - \frac{\lambda U}{W}, \quad c = \frac{\rho W \cos \alpha s U^2}{2(M+m)}, \quad d = \frac{\mu}{M+m}. \quad (70)$$

The stability criterion is as in (64); b must be negative, and

$$\begin{aligned} \frac{\mu}{M+m} &> \frac{\rho W \cos \alpha s^2 U^3}{16mg} - \frac{\lambda U}{2W} \\ &+ \frac{1}{2} \left(\left(\frac{\rho W \cos \alpha s^2 U^3}{8mg} - \frac{\lambda U}{W} \right)^2 - \frac{\rho^2 W^2 \cos^2 \alpha s U^5}{2mg(M+m) \left(\frac{\rho W \cos \alpha s^2 U^3}{8mg} - \frac{\lambda U}{W} \right)} \right)^{1/2}. \end{aligned} \quad (71)$$

Again we find the steady state is stable provided there is sufficient friction, and the critical level of friction increases with water speed. Above this, perturbations grow until they reach the water surface, and they then settle into a natural amplitude of bouncing. This amplitude is independent of the initial conditions - and the amplitude is larger for larger water speeds. Figure 22 shows example trajectories for U smaller than and larger than the threshold speed for instability, when the paddle is dropped into the water from a height. The same steady oscillations would be reached in the latter case if the paddle were initially perturbed slightly from the steady planing state.

6 Wagner model: potential flow

The previous sections have identified the rise of water ahead of the paddle, and the time lag necessary for this to happen, as a mechanism for generating super-elastic bounces from the water surface. We have attempted to model this in a shallow water setting and with an

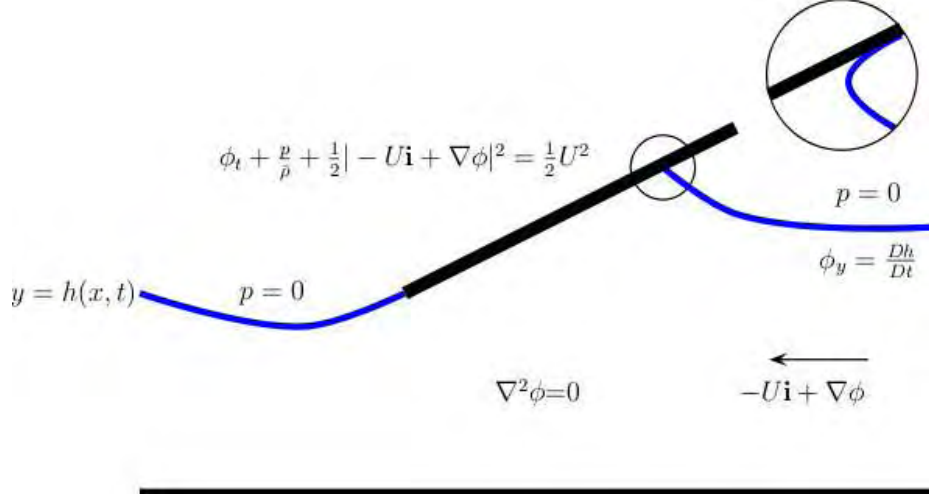


Figure 24: Potential flow problem for a paddle impacting the free surface of an inviscid irrotational fluid, moving at undisturbed speed U . The free surface is at atmospheric pressure $p = 0$, as well as being subject to a kinematic condition. The kinematic condition also holds on the paddle, and the pressure there determines the vertical force on the paddle, which can be used to determine how the paddle moves.

ad-hoc generalisation to deeper water. Both of these essentially use parameterisations for the effect of the paddle on the water flow.

A different approach is to attempt to calculate the water flow due to the paddle explicitly. This is again easiest in two dimensions, where we assume the flow is irrotational. We can then write the velocity as $\mathbf{u} = -U\mathbf{i} + \nabla\phi$, where ϕ is the disturbance potential satisfying Laplace's equation. We assume the stream occupies the space $-H < y < 0$, and consider the situation shown in figure 24. This type of problem has been reasonably well established as a so-called Wagner problem [20, 7, 5, 6].

The pressure is given by Bernoulli's equation, neglecting gravity,

$$\bar{\rho}\phi_t + p + \frac{1}{2}\bar{\rho}|-U\mathbf{i} + \nabla\phi|^2 = \frac{1}{2}\bar{\rho}U^2, \quad (72)$$

the constant being determined by a steady uniform upstream condition, and zero atmospheric pressure. The free surface $y = h(x, t)$ has kinematic condition $\phi_y = Dh/Dt$, and the pressure there must be atmospheric.

If a paddle enters the water, suppose it occupies the region of the surface $0 < x < L(t)$. L here is treated as the 'turnover' point (figure 24); there will be a small jet of water ejected ahead of the paddle, but the appreciable force comes only from the wetted region up to the turnover point L . For the purposes of this problem we disregard the details of this region and the high speed jet entirely; see [7] for further discussion. For this wetted part of the surface the relevant boundary condition is that the normal velocity of the water must equal the normal velocity of the paddle. If the paddle is a flat plate moving in a vertical plane with angle of attack α , then the normal is $\mathbf{n} = (-\sin\alpha, \cos\alpha)$, so that

$$-(-U + \phi_x)\sin\alpha + \phi_y\cos\alpha = \dot{z}\cos\alpha, \quad (73)$$

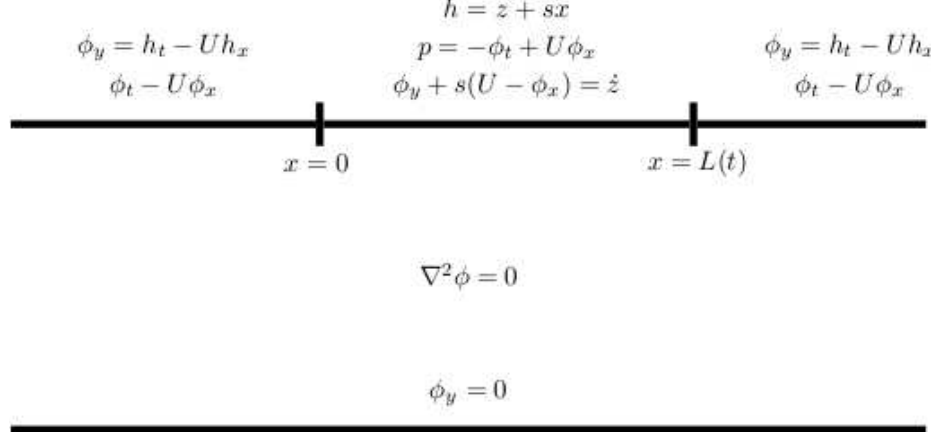


Figure 25: The linearised Wagner problem for a bouncing paddle $h_p = z + sx$ occupying the region $0 < x < L(t)$ of the surface of an inviscid irrotational fluid. The boundary conditions are as in figure 24, but are now applied at the undisturbed surface $y = 0$.

where \dot{z} is the vertical velocity of the paddle. We also know the height of the paddle

$$h_p(x, t) = z + x \tan \alpha, \quad (74)$$

where z is the vertical position of the trailing edge above the undisturbed surface $y = 0$. The unknown position of the leading edge $x = L(t)$ is determined by the condition that the free surface height here is equal to the height of the paddle:

$$h(L(t), t) = h_p(L(t), t). \quad (75)$$

If we consider the linearised form of this problem (figure 25) we have

$$\nabla^2 \phi = 0, \quad -H < y < 0; \quad \phi_y = 0 \quad \text{at} \quad y = -H; \quad \phi_x \rightarrow 0 \quad \text{as} \quad x \rightarrow \pm\infty, \quad (76)$$

At the surface $y = 0$,

$$\left. \begin{aligned} \phi_y &= h_t - U h_x, \\ 0 &= -\phi_t + U \phi_x, \end{aligned} \right\} \quad \text{for} \quad x < 0, \quad x > L(t), \quad (77)$$

$$\left. \begin{aligned} h &= h_p = z + sx, \\ \phi_y - s\phi_x &= \dot{z} - sU, \end{aligned} \right\} \quad \text{for} \quad 0 < x < L(t), \quad (78)$$

where $s = \tan \alpha$ is the slope of the paddle. At $x = L(t)$,

$$h(L(t), t) = h_p(L(t), t). \quad (79)$$

For a given $z(t)$ these provide the equations and boundary conditions to be solved for $\phi(x, y, t)$ and $h(x, t)$. However this is coupled to the equation of motion for the paddle,

$$(M + m)\ddot{z} = -mg + F_n \cos \alpha - \mu \dot{z} \quad (80)$$

since the normal force F_n is

$$F_n = \int_0^{L(t)} p \, dx = \bar{\rho} \int_0^{L(t)} -\phi_t + U\phi_x \, dx. \quad (81)$$

We can attempt to solve this problem numerically by solving Laplace's equation for ϕ at every time step using the pressure boundary conditions for $x < 0$ and $x > L$ and the kinematic boundary condition for $0 < x < L$ (the other boundary conditions are $\phi_y = 0$ at the bottom and $\phi_x = 0$ at the sides, which we try to make far away from the paddle). Then the kinematic boundary conditions on $x < 0$ and $x > L$ give the motion of the boundary there (with boundary condition $h = 0$ at $x = +\infty$), whilst Bernoulli's equation for $0 < x < L$ gives the pressure there and hence the acceleration of the paddle. We can therefore find the new position of the paddle and the free surface and find the new value of L at which these surfaces are equal. This gives the region over which the different boundary conditions should be applied for Laplace's equation at the next time step.

The solutions indicate that if the background stream flow U is large enough the paddle can receive a super-elastic kick from the water, just as we found in the shallow water model. In this case any initial condition will reach a steady bouncing state, with the amplitude dependent on the water speed and paddle mass. The water ahead of the paddle rises up to meet it, and the paddle does not leave the water until it is some distance above the undisturbed level at which it entered it. At slower speeds there seems to be a steady planing state in which the pressure balances the weight of the paddle.

It seems therefore that although the quantitative values (the amplitudes of the bounces and the speed at which bouncing starts) are different, the same instability mechanism is inherent in this model as in the shallow water model. This is reassuring since the potential flow captures more of the fluid dynamics than the shallow water assumptions do, but given the computational intensity, this model is not particularly useful nor wholly satisfying for understanding what is going on.

7 Conclusion

A light paddle towed along the surface of flat water will continually bounce, or skip, along provided it is pulled fast enough. The experiments we conducted (in which the paddle could move vertically and the water flowed past) suggest that for a given paddle mass there is a threshold water speed below which it planes steadily through the water and above which it oscillates up and down, leaving the surface and skipping if the oscillations are large enough. The amplitude of oscillations increases with the water speed above the threshold. For a given water speed there seems to be a natural bouncing mode (with characteristic amplitude and period) to which any initial condition evolves.

The threshold velocity at which skipping starts is larger for a heavier paddle and lower for a wider paddle, and for a given velocity lighter and wider paddles tend to bounce higher. When bouncing occurs the collision with the water appears to be super-elastic; the paddle gains energy from the water.

We have described several models to try to explain the experimental findings; the key mechanism we identify which causes the instability is the *transient* response of the water to

the paddle moving in and out. The most simple models which treat the water's response as a steady function of paddle displacement cannot explain the skipping. Instead, the variation of the 'wetted length' over the collision must be explicitly accounted for.

A shallow water description, in which the splashed up water ahead of the paddle is treated as a shock, indicates that there should be a bifurcation at a critical water speed above which the steady planing state becomes unstable and a bouncing mode, with super-elastic bounces, is reached. The amplitude of oscillations is larger for larger water speed and wider paddle, as found in the experiments.

A simplification and generalization to deeper water allow for a simple two o.d.e. description of the system which again has the bifurcation and correct dependence on the various parameters. The model is not quantitatively comparable to the experiment but this may be due to uncertainties in the parameterized effects of the water flow beneath and around the paddle.

A two dimensional linearized potential flow model verifies that accounting for the time evolution of the wetted length can allow super-elastic bounces and destabilize the steady planing state.

The shallow water model motivated a further experiment in which the paddle was placed in a shallow stream. We were able to verify the lag between vertical paddle position and wetted length during oscillations. This setup worked well and it would be worthwhile to do some further studies with varying the flow speed. The setup was more robust, easier to use, and avoids the complications of the wake in the rotating tank. With more careful measurements we might also be able to test more critically the models suggested here.

8 Acknowledgements

As well as my supervisor Neil Balmforth, to whom I owe a great deal of thanks for his never ending ideas and patience, I would also like to thank Jim McElwaine, whose enthusiasm and guidance set us off in the right direction at the beginning of the summer. I would also like to extend my thanks to all the many staff, fellows and visitors at Woods Hole who gave time to think about this problem on the porch at Walsh Cottage.

References

- [1] B. J. BINDER AND J.-M. VANDEN-BROECK, *Free surface flows past surfboards and sluice gates*, Eur. J. App. Math., 16 (2005), pp. 601–619.
- [2] L. BOCQUET, *The physics of stone skipping*, Am. J. Phys., 71 (2003), pp. 150–155.
- [3] C. CLANET, F. HERSEN, AND L. BOCQUET, *Secrets of successful stone-skipping*, Nature, 427 (2004).
- [4] A. E. GREEN, *The gliding of a plate on a stream of finite depth*, Proc. Camb. Philos. Soc., 31 (1935), pp. 589–603.
- [5] S. D. HOWISON, J. R. OCKENDON, AND J. M. OLIVER, *Deep- and shallow-water slamming at small and zero deadrise angles*, J. Eng. Math., 42 (2002), pp. 373–388.

- [6] ———, *Oblique slamming, planing and skimming*, J. Eng. Math., 48 (2004), pp. 321–337.
- [7] S. D. HOWISON, J. R. OCKENDON, AND S. K. WILSON, *Incompressible water-entry problems at small deadrise angles*, J. Fluid. Mech., 222 (1991), pp. 215–230.
- [8] W. JOHNSON, *The ricochet of spinning and non-spinning spherical projectiles, mainly from water. Part II: An outline of theory and warlike applications*, Int. J. Impact Engng., 21 (1998), pp. 25–34.
- [9] A. KOROBKIN, *Shallow-water impact problems*, J. Eng. Math., 35 (1999), pp. 233–250.
- [10] H. KUNINAKA AND H. HAYAKAWA, *Anomalous behavior of the coefficient of normal restitution in oblique impact*, Phys. Rev. Lett., 93 (2004).
- [11] M. Y. LOUGE AND M. E. ADAMS, *Anomalous behavior of normal kinematic restitution in the oblique impacts of hard sphere on an elastoplastic plate*, Phys. Rev. E, 65 (2002).
- [12] P. R. PAYNE, *The water rise in front of a model planing hull*, Experiments in Fluids, 17 (1994), pp. 96–104.
- [13] L. ROSELLINI, F. HERSEN, C. CLANET, AND L. BOCQUET, *Skipping stones*, J. Fluid. Mech., 543 (2005), pp. 137–146.
- [14] T. SUGIMOTO, *Mechanics of the surf skimmer revisited*, Am. J. Phys., 71 (2003), pp. 144–149.
- [15] N. TABERLET, S. W. MORRIS, AND J. N. MCELWAIN, *Washboard road: The dynamics of granular ripples formed by rolling wheels*, Phys. Rev. Lett., 99 (2007).
- [16] L. TING AND J. B. KELLER, *Planing of a flat plate at high Froude number*, Phys. Fluids., 17 (1974), pp. 1080–1086.
- [17] E. O. TUCK AND A. DIXON, *Surf-skimmer planing hydrodynamics*, J. Fluid. Mech., 205 (1989), pp. 581–592.
- [18] J.-M. VANDEN-BROECK, *Numerical calculations of the free-surface flow under a sluice gate*, J. Fluid. Mech., 330 (1997), pp. 339–347.
- [19] T. VON KARMAN, *The impact of sea planes during landing.*, NACA Tech. Note 321, (1929).
- [20] H. WAGNER, *Über Stoß- und gleitvorgänge an der Oberfläche von Flüssigkeiten (Phenomena associated with impacts and sliding on liquid surfaces)*, Zeitschr. Angew. Math. Mech., 12 (1932), pp. 193–215.

On the Organization of Geostrophic Circulations Over Large Scale Topography by Eddy-Diffusion of PV

Malte Jansen

January 12, 2009

1 Introduction

The quasi-geostrophic (QG) potential vorticity (PV) for a homogeneous layer can be written as

$$q = \beta y + \zeta - f \frac{h'}{H} \quad (1)$$

where β is the planetary PV gradient ($\frac{\partial f}{\partial y}$), ζ is the relative vorticity, H is the mean depth of the layer and h' is a perturbation of the latter. For a layer bounded below by the bottom topography we can write:

$$q = \beta y + \zeta - f \frac{\eta}{H} + f \frac{h_b}{H}, \quad (2)$$

where η is the displacement of the interface bounding the layer at the top and h_b is the elevation of the bottom topography. A topographic feature can thus be viewed as an anomaly in the in the background PV field, similar to the β -effect.

Different theoretical concepts have been proposed arguing that quasi-geostrophic turbulence acts to "homogenize" the mean potential vorticity (PV) field which can be "disturbed" by topographic features. (See [7] for an older but still very useful review.) This homogenization would then lead to an anomaly in the relative PV field¹ over the topography which would be associated with a mean circulation.

Indeed observational evidence exists for this kind of circulations around topographic features. [8] first noticed an almost barotropic 100Sv anticyclonic flow around the Zapiola Drift in the South Atlantic. A result later confirmed by various other observational studies.

Many previous theoretical and numerical studies focused on the adjustment of a turbulent flow in a closed domain in the absence of forcing and with weak or no dissipation (e.g. [3] or [5]). [2], however, suggests that the strength of topographic circulations might be determined by a balance between the eddy flux of PV, and Ekman pumping due to bottom friction. Kinetic energy is in this case assumed to be provided by eddies from remote regions (such as the ACC for the Zapiola drift). The same approach is used in this study.

¹relative PV is here used as the PV anomaly given by the displacement of the free interface and the relative vorticity

A simple qualitative explanation for the tendency of turbulent flows to homogenize PV can be given by assuming that stirring by geostrophic turbulence moves elements of fluid over the topography. Assuming that this happens on time-scales shorter than those of forcing and dissipation, these fluid elements will conserve their PV. If the fluid is stirred randomly so that each fluid parcel reaches each point of a domain with a similar average frequency, the time-mean PV has to be constant over that domain. If we further assume that the stirring "transports" inhomogeneities to smaller scales, until viscosity finally mixes the fluid irreversibly, the PV field will be homogenized.

Here we want to follow an approach used by [2] which is based on this idea of eddy-mixing. Treating PV as a conserved passive tracer we apply the commonly used approach of down-gradient eddy-diffusion. It should be noted that in doing this we ignore the fact that PV is not strictly a passive tracer since it's distribution determines the flow field. However, the results of eddy resolving numerical simulations, presented in this study, tend to support the use of this approximation.

An analytical derivation for the expected mean flow over topography in a two layer QG model assuming down-gradient diffusion of PV is shown in section 2. In section 3 the predictions of this theory are tested against eddy-resolving numerical simulations and section 4 presents some preliminary results from laboratory experiments. A concluding discussion of the results is given in section 5.

2 A Two Layer QG Model Assuming Down-Gradient Diffusion of PV

The two layer quasi-geostrophic PV equations on an f-plane with horizontal viscosity ν and linear bottom friction $F = -Ru$, can be written as

$$\left[\frac{\partial}{\partial t} + J(\psi_1, \cdot) \right] q_1 = \nu \nabla^4 \psi_1 \quad (3)$$

$$\left[\frac{\partial}{\partial t} + J(\psi_2, \cdot) \right] q_2 = -R \nabla^2 \psi_2 + \nu \nabla^4 \psi_2 \quad (4)$$

where

$$q_1 = \nabla^2 \psi + \frac{f}{H_1} \eta \quad (5)$$

$$q_2 = \nabla^2 \psi + \frac{f}{H_2} (h_b - \eta). \quad (6)$$

The shear between the two layers is given by the interface displacement η through the thermal wind relation

$$\psi_1 - \psi_2 = \frac{g'}{f} \eta. \quad (7)$$

Assuming that the eddy flux of PV can be described by down-gradient diffusion

$$\overline{u'q'} = -K \frac{\partial \bar{q}}{\partial x}, \quad (8)$$

the steady-state solution would be given by:

$$(K_1 + \nu)\nabla^4\psi_1 + \frac{K_1 f}{H_1}\nabla^2\eta = 0 \quad (9)$$

$$(K_2 + \nu)\nabla^4\psi_2 + \frac{K_2 f}{H_2}\nabla^2(h_b - \eta) - R\nabla^2\psi_2 = 0, \quad (10)$$

where where the eddy diffusivity of PV, $K_{1/2}$ is allowed to be different in the two layers. If h_b , η and ψ_2 vanish somewhere outside the domain of interest, (9) gives:

$$\zeta_1 \equiv \nabla^2\psi_1 = -\frac{K_1}{K_1 + \nu}\frac{f}{H_1}\eta. \quad (11)$$

Using (7), we find

$$\left(\frac{K_1 + \nu}{K_1}L_{D_1}^2\nabla^2 + 1\right)\psi_1 = \psi_2 \quad (12)$$

where $L_{D_1}^2 = \frac{H_1 g'}{f^2}$. Hence we find a barotropic solution ($\psi_1 \approx \psi_2$) if

$$\frac{L_{D_1}^2}{L^2} \ll \frac{K_1}{K_1 + \nu}, \quad (13)$$

where L is the typical length-scale of the topographic mean flow which is assumed to be of the same scale as the topography itself. Note that the left hand side corresponds to the Burger number. With (13), we find from (11) that

$$\eta \approx -\frac{K_1 + \nu}{K_1}\frac{H_1}{f}\nabla^2\psi_2 \quad (14)$$

and therefore

$$\frac{\eta}{H_1} \sim \frac{K_1 + \nu}{K_1}Ro \quad (15)$$

where

$$Ro = \frac{U}{fL}. \quad (16)$$

We now find that (10) simplifies to

$$\psi_1 = \psi_2 = \frac{K_2 f}{RH_2}h, \quad (17)$$

if

$$\frac{L_{D_1}^2}{L^2} \ll \frac{K_1}{K_1 + \nu}, \quad Ro \ll \frac{K_1}{K_1 + \nu}\frac{h}{H_1} \quad \text{and} \quad Ro \ll \frac{K_2}{K_2 + \nu}\frac{h}{H_2}. \quad (18)$$

Note that the first two conditions imply that $\eta \ll h$.

For sufficiently small Rossby and Burger numbers, we find a barotropic flow along isolines of constant depth controlled by an equilibrium between an eddy-diffusive flux of PV and bottom friction. This result is similar to the result obtained by [2] in the limit of a vanishing planetary vorticity gradient and wind-stress.

3 Numerical Simulations

The full set of equations (3) to (7) are integrated numerically. A list of parameters is given in table 3, they are chosen to resemble the laboratory experiments described later in this report. The topography used is given by a truncated cone centered in the middle of the domain. The length scale L_{top} given in table 3 denotes the length of the slope, the total diameter of the "bump" is $3L_{top}$. It should be noted that the topography is large compared to the deformation radius. Additionally to the truncated cone the topography is slightly parabolic (maximum amplitude about $0.1 H_2$) over the whole domain to simulate the parabola arising due to the centrifugal force in the tank.

L_{domain}	L_{top}	L_D	H_1/H_2	h_b/H_2	f	R	ν
150 cm	17 cm	1.5 – 2 cm	0.5	0.5	$3s^{-1}$	$0.01 - 0.08 s^{-1}$	$0.01 cm^2 s^{-1}$

Table 1: Parameters used for the numerical simulations

The equations are intergrated using a pseudo-spectral² model with a grid spacing $\Delta \approx 0.6cm$. Considering that the typical eddy scale is a few times the deformation radius the model is eddy resolving. The model has periodic boundary conditions though a sponge layer is implemented along the boundaries which strongly damps PV perturbations and thus eliminates eddies that move into this boundary layer. The sponge layer was chosen to simulate an infinite domain in which eddies are allowed to move out of the area of interest. Turbulence was generated by four pairs of sources and sinks of PV located in the middle of the four sides of the square domain, right outside the sponge layer.

For the topography used in the simulation (17) gives a constant along-slope velocity

$$u_\phi = \frac{\partial \psi}{\partial r} = \frac{K_2 f}{R H_2} \frac{\partial h_b}{\partial r} = -\frac{K_2}{R} \frac{h_b f}{H_2 L_{top}} \equiv -\frac{K_2}{R} \beta, \quad (19)$$

where we defined β as the topographic PV gradient. Nondimensionalizing all variables using the deformation radius L_D as a length scale and $L_D^{-1} \beta^{-1}$ as a time scale gives

$$u_\phi^* = -\frac{K_2^*}{R^*}. \quad (20)$$

Two series of simulations are performed in which first the strength of the forcing was varied, and then the magnitude of bottom friction. Each simulation is run into a statistically steady state for at least 500 s. After this initialization period the PV fields of each layer are averaged over at least 1000 s. An overview of all simulations is given in table 2. Figure 1 shows an example of mean relative PV and streamfunction fields from the S04-3 simulation. We see the four PV sources which produce jets that get deflected to the right, which can be explained by the parabolic background PV field. Over the topography we observe anticyclonic circulations along the slope in both layers which is in qualitative agreement with (19). This is associated with negative relative PV in the lower layer and a positive relative PV in the upper layer, which is the PV disturbance expected from a positive displacement of

²Pseudo spectral here means that the product $\mathbf{u}q$ in the nonlinear advection term is calculated in grid-space and then transformed back into fourier space

the interface. The negative relative PV in the lower layer acts to reduce the topographic PV anomaly, though it does not completely compensate the latter which is as high as $\frac{fh_b}{H_2} \approx 1.7$. Moreover we can find that a strong PV anomaly is only found on the lower part of the slope with a relatively sharp transition to a smaller PV anomaly further inside. This indicates a "PV staircase" as suggested by [7] due to a limitation of the cross-slope eddy mixing length to the Rhines-scale ($L_R = \sqrt{u/\beta}$) which is of the same order as the deformation radius here.

Name	S01	S02	S04-1	S04-2	S04-3	S08	S16	S32	R1	R2	R4	R8
Source [s^{-2}]	0.1	0.2	0.4	0.4	0.4	0.8	1.6	3.2	0.4	0.4	0.4	0.4
R [$10^{-2}s^{-1}$]	2	2	2	2	2	2	2	2	1	2	4	8

Table 2: List of performed numerical simulations. The second line gives the peak absolute value of the vorticity sources and sinks and the last line gives the bottom drag. The deformation radius L_D was 2 cm for simulations S01-S32 and 1.5 cm for simulations R1-R8. All other parameters are as given in table 1.

For a more quantitative analysis the mean along slope velocity was calculated for each performed simulation. Figure 2 shows the non-dimensional mean along-slope velocity from simulations S01 to S32 against the non-dimensional mean lower layer Eddy Kinetic Energy (EKE) averaged over the whole domain of integration and against the EKE averaged only over the topography. A first observation is that the along slope velocity increases with EKE. Since we expect the eddy diffusivity to rise with EKE, this is in qualitative agreement with the theoretical prediction (20). However, while the flow is anticyclonic in both layers, the magnitude is quite different in some simulations, especially for those with low EKE where we find a stronger anticyclone in the upper layer than in the lower layer. As the strength of the sources is increased and the EKE becomes higher, the flow becomes more barotropic with the anticyclone in the lower layer being somewhat stronger than in the upper layer for high EKE. Focussing on the lower layer velocities and the EKE over the topography (which we assume is a better predictor for the eddy dissusivity in this area) we find that the along-slope velocity approximately grows as $EKE^{3/4}$.

The left panel in figure 3 shows the along-slope velocity versus the inverse of the bottom friction for experiments R1 to R8. The linear increase of u_ϕ^* with $1/R^*$ appears to be in good agreement with (20), assuming that the eddy-diffusivity does not change between the experiments. This, however, can not necessarily be expected.

To allow for a better test of prediction (20), two different methods were used to estimate the eddy-diffusivity K_2 and how it relates to EKE.

The eddy-flux of PV can generally be decomposed into a down-gradient flux and a bolus flux along iso-lines of PV as follows:

$$\overline{\mathbf{u}'q'} = \frac{\overline{\mathbf{u}'q'} \cdot \nabla \bar{q}}{|\nabla \bar{q}|^2} \nabla \bar{q} - \frac{\overline{\mathbf{u}'q'} \times \nabla \bar{q}}{|\nabla \bar{q}|^2} \times \nabla \bar{q} \quad (21)$$

$$\equiv -K \nabla \bar{q} + \psi_B \times \nabla \bar{q}. \quad (22)$$

Using this formulation we can calculate the actual PV-diffusivity K and bolus stream-function ψ_B from time averages of the numerical integration. While ψ_B is rather indistin-

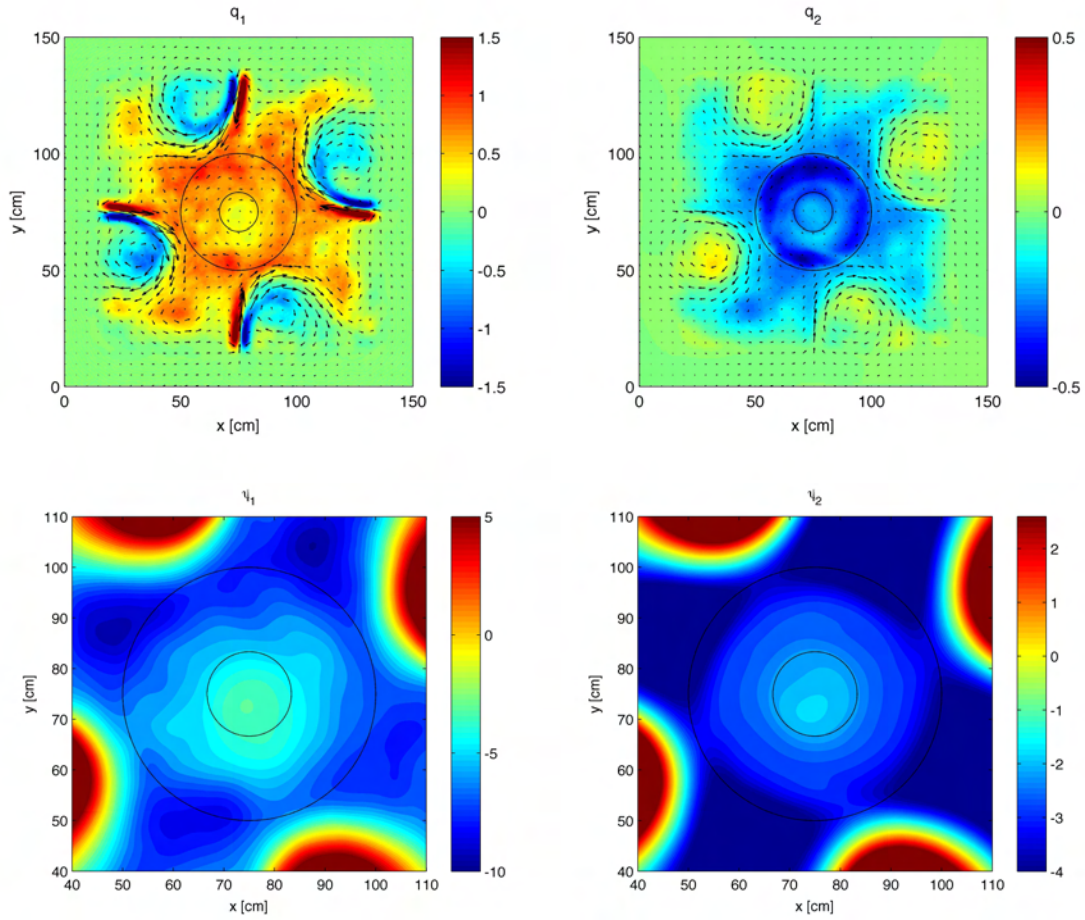


Figure 1: Mean relative PV in s^{-1} and velocity vectors (top) and streamfunction in $cm^2 s^{-1}$ (bottom) for simulation S04-3. The upper layer is shown on the left and the lower layer is shown on the right. The black circles indicate the topography, with the outer circle representing the outer edge of the cone and the inner circle representing the truncated top.

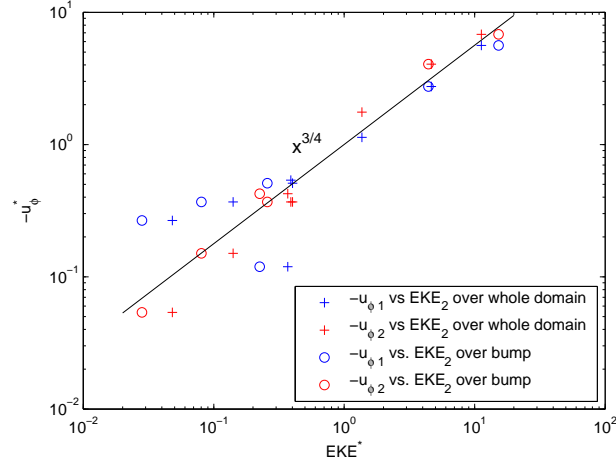


Figure 2: Non-dimensional mean along-slope velocity u_ϕ in the upper (blue) and lower layer (red) for simulations S01 to S32 vs. the (non-dimensional) mean EKE averaged over the whole domain of integration (crosses) and against the EKE averaged over the topography (circles).

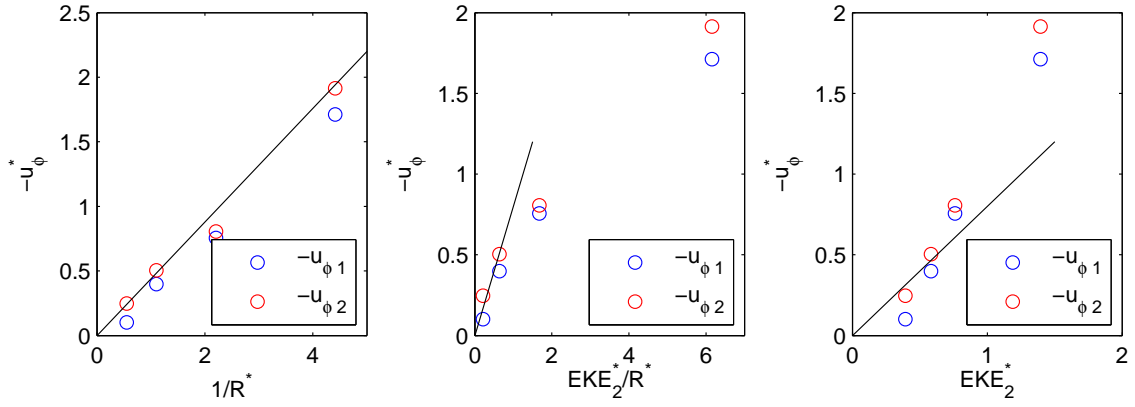


Figure 3: **Left:** Non-dimensional mean along-slope velocity u_ϕ in the upper (blue) and lower layer (red) for simulations R1 to R8 against the inverse of the non-dimensional friction parameter ($1/R^*$). **Center:** As left, but u_ϕ against EKE_2^*/R^* , where EKE^* denotes EKE averaged over the topography. **Right:** As left but u_ϕ against EKE_2^*

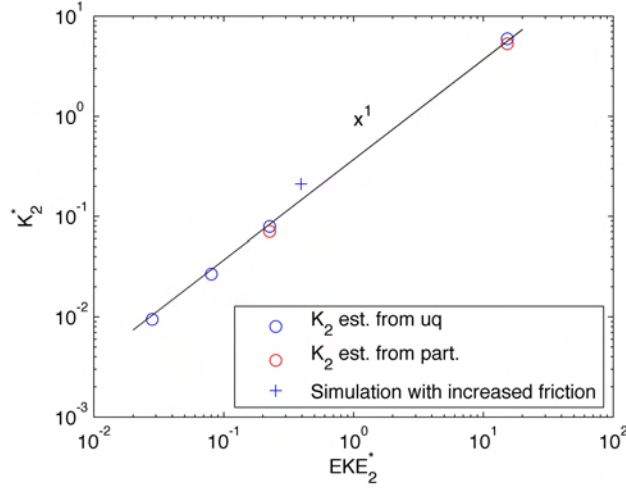


Figure 4: Lower level eddy-diffusivity K_2^* over the topographic slope, estimated from (22) (blue) and from (23) (red) against EKE_2^* over the topography. The circles show data from experiments S01, S02, S04-3 and S32. The cross shows an estimate from R8.

guishable from noise in both layers, for the data available, the lower layer diffusivity K_2 is found to be coherently negative over the slope of the topography, which retrospectively justifies the down-gradient diffusion parameterization used in the theory presented above. An averaged eddy-diffusivity over the slope is calculated this way.

This direct calculation of PV diffusion is compared to the diffusion of a conserved tracer obtained from particle release simulations. For two experiments (S04-3 and S32) 1000 particles were placed in a circle along the topographic slope. The diffusivity can be estimated using

$$4K = \frac{\partial}{\partial t} \left\langle \frac{1}{N} \sum_i r_i^2 \right\rangle, \quad (23)$$

where N is the total number of particles and r_i is the distance of each particle from the center. Since we are interested in estimating the diffusivity above the slope, we evaluate (23) only until the particles leave the slope-area. To increase the accuracy of the estimate, the tracer release was repeated 5 and 8 times in succession for the S04-3 simulation and the S32 simulation, respectively. A possible error in this estimate up to a factor of two however, should still be considered.

Figure 4 shows the eddy-diffusivity over the slope in the lower layer, estimated in this two ways versus the lower layer EKE over the topography. Note, that for the two experiments for which both methods were applied, the eddy diffusions of PV calculated directly from (22) agree well with the diffusivity estimates from the particle release simulations. We further find that the eddy diffusivity scales convincingly like

$$K^* \sim EKE^*, \quad (24)$$

though simulation R8 for which Ekman friction was increased by a factor of four deviates somewhat from this scaling. This result might come rather surprising considering that for

a fully turbulent flow we expect a scaling like

$$K \sim U_{eddy} L_{eddy} \quad (25)$$

with typical eddy velocities and scales U_{eddy} and L_{eddy} , respectively. Assuming that L_{eddy} scales like the deformation radius this becomes, in non-dimensional parameters

$$K^* \sim U_{eddy}^* \sim \sqrt{EKE^*}. \quad (26)$$

This is not observed, probably because flow over the slope is not fully turbulent. Comparing the nonlinear advection term of the QG PV equation with the advection of topographic PV (the topographic β -term) we find a “nonlinearity -parameter”³

$$N \equiv \frac{J(\psi, \nabla^2 \psi)}{\frac{1}{r} \frac{\partial \psi}{\partial \phi} \beta} \sim \frac{L_R^2}{L_D^2} = U_{eddy}^*, \quad (27)$$

where $L_R = \sqrt{\frac{U_{eddy}}{\beta}}$ is the topographic Rhines scale. N ranges around 0.2 to 5 in the simulations shown here, indicating that the assumption of fully developed turbulence is rather inadequate. Considering that (undamped) linear motion does not contribute to mixing we can argue that the eddy-diffusion in this weakly nonlinear case is given by the traditional scaling for turbulent motion (25) times this nonlinearity-parameter which would give

$$K^* \sim U_{eddy}^* N \sim EKE^*. \quad (28)$$

However, to some extent any power of N and thus any power of U^* could be justified this way. Since the eddy scale is generally larger than the deformation radius, which makes the effective nonlinearity even smaller than the above estimate, we could argue that the motion over the slope is in a regime of linear damped waves. In this case the “eddy diffusivity” would be given by

$$K = \int u(t)u(t+\tau)d\tau \sim \frac{U^2 R}{R^2 + \omega^2} \approx \frac{U^2 R}{\omega^2} \quad (29)$$

where ω is the frequency, U is the velocity perturbation associated with the wave motion and R is the coefficient of bottom friction (see above). The approximation on the right-hand side holds for weakly damped waves. Assuming that the wave length scales like the deformation radius and the corresponding frequency is given by the Rossby-wave dispersion relation we find, in non-dimensional parameters

$$K^* \sim \frac{U^{*2} R^*}{R^{*2} + 1} \approx U^{*2} R^* = EKE^* R^*. \quad (30)$$

This limit can explain the observed linear scaling of K^* with EKE^* . However, the calculated eddy diffusivity for simulation R8 agrees somewhat less with (30) than with (28). Plotting K^* vs. $EKE^* R^*$ shows that the diffusivity of R8 lies significantly below the line $K^* \sim EKE^* R^*$ (not shown).

The center and right panels of figure 3 show the along slope velocity versus EKE^*/R^* which would be expected to be linear from (20) and assuming the scaling law (28), and versus

³The inverse of this parameter is sometimes referred to as “steepness”.

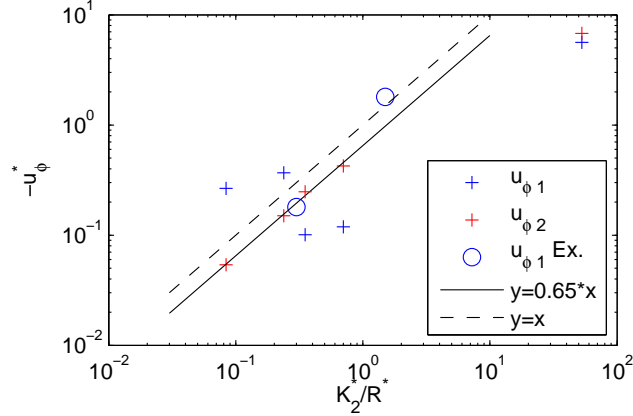


Figure 5: Non-dimensional along-slope velocity u_{ϕ}^* in the upper (blue) and lower layer (red) versus K_2^*/R^* for all simulations for which K_2 was explicitly calculated (S01, S02, S04-3, S32 and R8). The circles show two experimental results (see text). The solid straight line denotes $-u_{\phi}^* = 0.65 \cdot K_2^*/R$ and the dashed line shows $-u_{\phi}^* = K_2^*/R$ which is the theoretical result (20).

EKE^* which would be expected to be linear assuming scaling law (30), for simulations R1 to R8. Both relations appear rather non-linear, suggesting that both scaling laws might be inadequate to describe the scaling of the eddy diffusivity under variation of bottom friction, found here.

The most direct test of the theoretical prediction (20) is given by figure 5 which shows the along-slope velocity in both layers versus K_2^*/R^* for all simulations for which K_2 was explicitly calculated. As mentioned before, we note the lack of barotropization especially for low-EKE simulations, which was expected from the theory. For the lower layer, however, we find good agreement with the theory as long as the EKE does not become too large. For high EKE (i.e. high eddy diffusivity) the results indicate a deviation towards lower along-slope velocities than expected from (20).

4 Experimental Results

In a second step the theory shall be verified by laboratory experiments. Experiments were conducted in a rotating tank using a two layer fluid. The density difference was obtained using water with slightly different salinity. The reduced gravity for the experiment shown here is $g' = 4 \text{ cm s}^{-2}$. The rotation rate gives $f = 4 \text{ s}^{-1}$, which results in a deformation radius $L_D \approx 2 \text{ cm}$. A truncated cone is placed in the center of the tank as in the numerical simulations. A sketch of the experiment is shown in figure 6. Eddies are generated by four pairs of sources and sinks in the upper layer equally spaced around the tank. The method is described in detail in [4].

Since the expected along-slope velocities are on the order of 1 mm s^{-1} or less the experiment, and especially the filling up of the upper layer, has to be done very carefully. A series of similar experiment has previously been performed by [6], but the results are not used

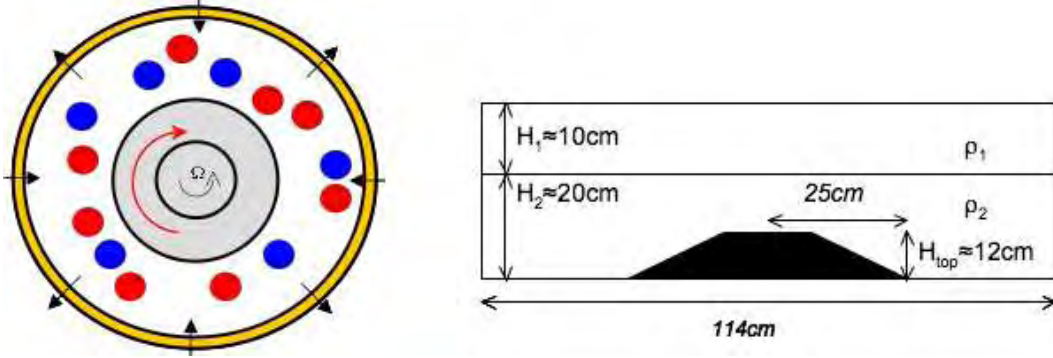


Figure 6: Sketch of the experimental setup. Left: Topview. The arrows on the tank wall indicate sources and sinks in the upper layer to generate pairs of cyclonic and anticyclonic eddies, indicated in red and blue. Right: Side view. The figure on the left is taken from [1].

here, mainly because a review of the lab-recordings suggested that the filling-up procedure in these experiments might have had a significant influence on the observed circulation, which can hardly be filtered out in hindsight. Due to similar and other problems during this study, only the results of one experiment will be discussed here. However, the flow rate of the sources and sinks was increased by a factor of 1.7 during this experiment, so that two different values of the flow rate are presented. The first part of the experiment, with a flow rate of 540 cc/min per jet, will be referred to as Ex3, while the second part, with a flow rate of 920 cc/min per jet, will be referred to as Ex3.5, in the following. Since the diffusion of the interface between the two layers is significantly enhanced due to the eddy activity, it is likely that a sharp interface did not exist any more during Ex3.5, which therefore has to be handled with care.

Figure 7 shows a series of snapshots from Ex3. Neutrally buoyant dye, released from outlets on the bottom of the cone, indicates an anticyclonic circulation around the cone in the lower layer, with an average velocity of the order of $-u_{\phi 2} \approx 0.3 \text{ mm/s}$. The flow in the upper layer can be evaluated more quantitatively by tracking of surface particles which can also be seen in figure 7. Two 30-minute time averages of surface velocity and vorticity from Ex3 are shown in figure 8. The most striking feature that is observed is a strong cyclone above the flat center of the truncated cone. It should be noted that the frame-rate used for the particle tracking was slower for the figure shown on the left, which causes an inadequate resolution of the fast cyclone observed over the top of the cone. The cyclone remained above the flat center of the cone during the whole time of the experiment. Dye releases in the lower layer indicate an associated cyclonic motion in the lower layer, though probably weaker. A similar feature is not predicted by theory and was never observed in the numerical simulations. It should be noted that the Rossby-number for this cyclone is $O(1)$, which might explain its absence in the numerical QG model. A physical explanation for this observation, however, is still outstanding. Over the slope itself we observe an anticyclonic circulation as predicted by the theory. The averaged along-slope velocity in the upper layer is $-u_{\phi 1} \approx 0.9 \text{ mm/s}$ and thus probably stronger than in the lower layer, similar to the results found in the numerical simulation for rather small EKE. We further

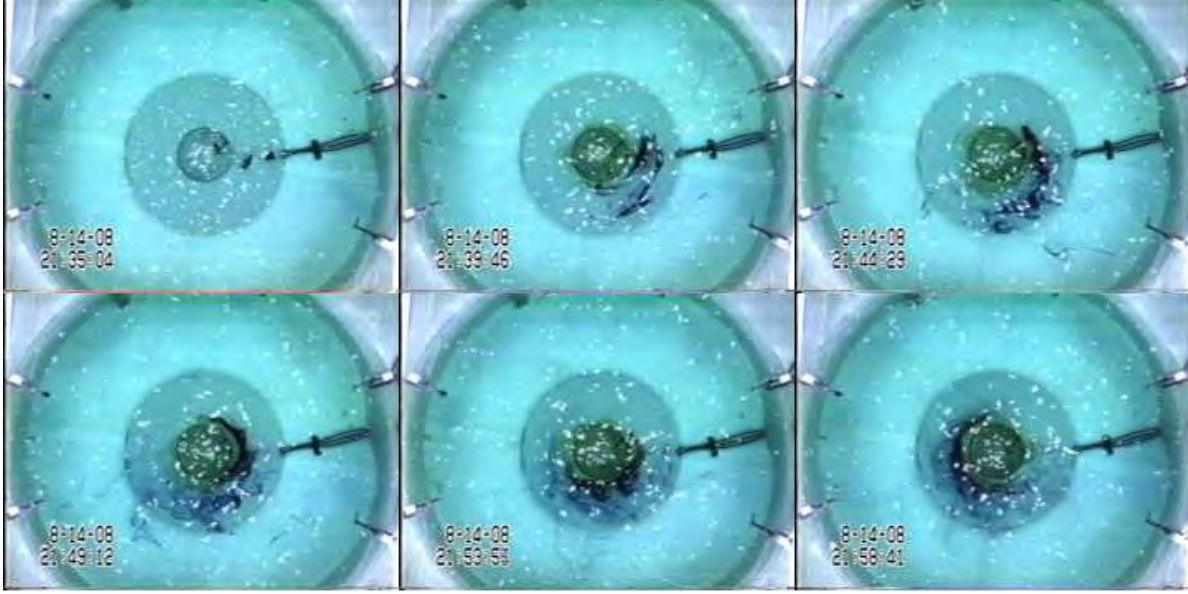


Figure 7: Series of snapshots from Ex3. Time elapses from the top left to the bottom right with the time difference between each snap-shot being about 4min 45s. One can see the dye in the lower layer traveling anticyclonically along the topographic slope.

observe a band-like vorticity structure. This was previously mentioned in the discussion of the numerical results and probably indicates a PV-staircase as discussed by [7].

Figure 9 shows a ten minute average⁴ of surface velocity and vorticity from Ex3.5, i.e. with increased flow rate. Since the kinetic energy of the fluid injected by the sources scales with the square of the flow-rate, and the amount of fluid with that kinetic energy scales linearly with the flow rate, we expect the kinetic energy injected per time to scale like the cube of the flow rate. Assuming linear dissipation, we thus expect the EKE in the tank to rise approximately like the cube of the flow rate. An increase of the flow rate by a factor of 1.7 should therefore cause an increase of the EKE by a factor of $1.7^3 \approx 5$. Using the scaling (28) or (30), the theoretical result (20) would therefore suggest an increase of the along-slope velocity by a factor of 5. The experiments show an increase of the mean along-slope velocity by about a factor of ten. The estimated upper layer mean velocity over the slope for Ex3.5 is $-u_{\phi 1} \approx 9 \text{ mm/s}$.

The results of experiments Ex3 and Ex3.5 are added to figure 5. Since no good estimate for the eddy diffusivity exists for the experiments exists, for Ex3 K_2^* is fitted to match with the numerical results. For Ex3.5 K_2^* was then chosen to be five times this value, following the argumentation above. Due to the "tuning" of the diffusivity of Ex3, the comparison to the numerical result has some limitations. However, figure 5 indicates that the experiments are in a regime well covered by the numerical study. Also the slope between the two experimental data points, which is not affected by this tuning, shows general agreement with the numerical results. Note, that the used EKE estimate provides an estimate for EKE

⁴Since the faster flow allows for a reasonable average in a shorter time and the strong eddy activity in Ex3.5 caused the interface to be destroyed more rapidly, a shorter time average is used in Ex3.5 than in Ex3.

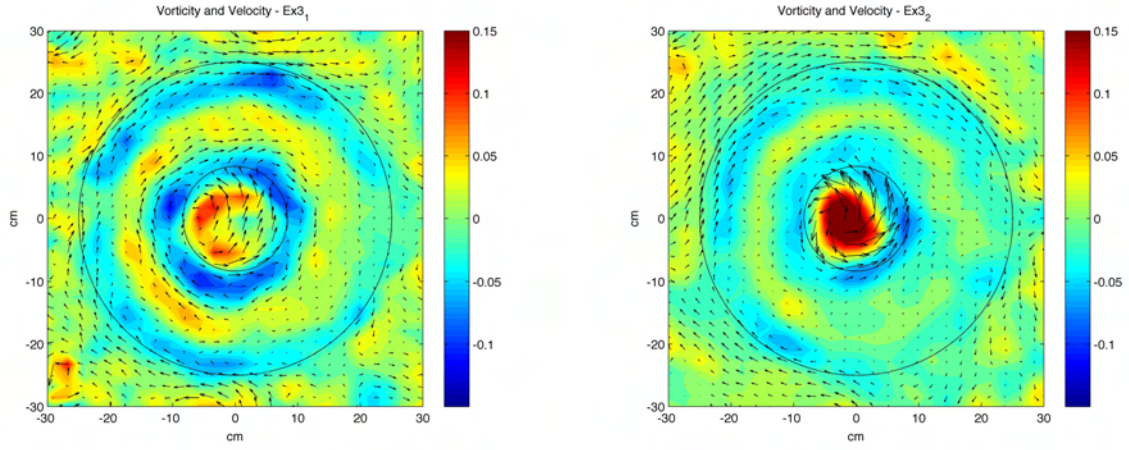


Figure 8: Two different eulerian mean velocity fields and the resulting vorticity, each averaged over about 30 minutes of EX3. The average field shown on the left was taken about one hour earlier than the field on the right. Velocity vectors are stretched by a factor of 10. The solid lines indicate the truncated cone.

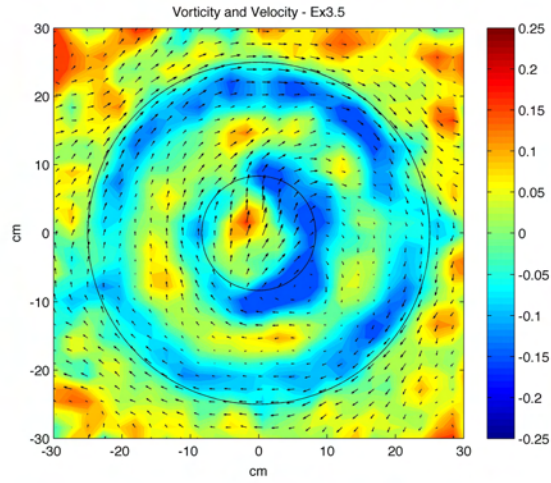


Figure 9: Eulerian mean velocity fields and the resulting vorticity each averaged over about 10 minutes of EX3.5. Velocity vectors are stretched by a factor of 2

averaged over the whole domain and not over the topography itself. The numerical models showed that the latter might grow faster with increased forcing, since the topographic PV gradient prevents weak eddies to propagate over the topography. Note also, that, due to the better measurability, the experimental results show upper layer velocities. As in the numerical simulations, however, the flow was not strictly barotropic.

5 Summary and Discussion

Assuming down-gradient eddy-diffusion of potential vorticity it was shown that in a two layer QG model for small Rossby numbers an approximately barotropic mean along-slope circulation is expected to arise over an isolated topographic feature which is anticyclonic for a mount and cyclonic for a depression. The strength of the circulation is controlled by an equilibrium between the eddy flux of PV and bottom friction. This prediction was tested using numerical simulations performed with an eddy resolving two-layer QG model and laboratory experiments.

The numerical simulations reproduce an anticyclonic along-slope circulation in both layers over a mount. However, especially for low EKE, the magnitude of the circulation is significantly different in the two layers, which disagrees with the theoretical prediction of a barotropic flow. Part of this inconsistency might be explained by a direct influence of the forcing. The latter consists of four pairs of PV sources and sinks in the upper layer which generate jets penetrating towards the center of the tank. The jets get unstable and generate eddies. These jets, however, could have a significant direct influence on the circulation in the upper layer, especially in the less turbulent simulations with low EKE.

The eddy diffusivity was estimated for a series of experiments. The results suggest a down-gradient diffusion of PV over the topography of similar magnitude as the eddy diffusivity of a passive tracer. It was further found that in the performed simulations the eddy-diffusivity over the slope scales linearly with the EKE, which suggests a linear or weakly nonlinear motion over the slope. However, no scaling law was found which consistently describes the dependence of the observed eddy diffusivity on bottom friction. It is likely that the experiments are covering a transition between a regime of fully linear damped waves and weakly nonlinear motion, for which case the dependence on bottom friction is expected to change over the range of performed experiments.

In the simulations for which the eddy diffusivity is calculated explicitly, the lower layer velocity agrees well with the theory in the range of lower and medium eddy diffusivity, but the along-slope velocity is smaller than the theoretical prediction for high eddy diffusivity.

It seems desirable to test whether the differences between the eddy resolving numerical simulation and the theoretical results, are mainly a shortcoming of the down-gradient diffusion parameterization or due to additional approximations made for the analytical solution. Preliminary results, not shown above, solving the full two layer QG model numerical without forcing but using a down-gradient diffusion parameterization, indicate that the differences are mostly a shortcoming of the down-gradient diffusion parameterization (including the lack of a representation of the direct effect of the forcing) and not dominantly due to inappropriate additional approximations made to obtain a simple analytical solution. Adding an additional eddy-viscosity seems to improve the results, though no convincing physical justification for this is known to the author.

Laboratory experiments using a two-layer fluid in a rotating tank do also reveal an anticyclonic flow along the topographic slope, which is in qualitative and to a lesser extend quantitative agreement with the theory. However, additionally a strong cyclone was observed over the flat center of the topography. An explanation for this observation is so far outstanding. More, carefully performed, experiments would be necessary to test the theoretical predictions in a more quantitative way.

6 Acknowledgments

I would like to thank Claudia Cenedese for her support during the whole project. Thanks to Bill Dewar for sparking my interest in this problem and to Raffaele Ferrari and John Whitehead for fruitful discussions. Special thanks to my fellow fellows for many great scientific discussions and a good time.

References

- [1] C. CENEDESE AND W. K. DEWAR, *Collaborative research: The zapiola blender*. Project Proposal submitted to NSF.
- [2] W. K. DEWAR, *Topography and barotropic transport control by bottom friction*, J. Mar. Res., 56 (1998), pp. 295–328.
- [3] G. HOLLOWAY, *Systematic forcing of large-scale geophysical flows by eddy-topography interaction*, J. Fluid Mech., 184 (1987), pp. 463–476.
- [4] P. F. LINDEN, B. M. BOUBNOV, AND S. B. DALZIEL, *Source–sink turbulence in a rotating stratified fluid.*, J. Fluid Mech., 298 (1995), pp. 81–112.
- [5] D. P. MARSHALL, S. T. ADCOCK, AND C. E. TANSLEY, *Geostrophic eddies, abyssal recirculations, and zonal jets*, 2001. Conference paper.
- [6] R. B. D. MESQUITA, *Experimental study of bottom controlled flow over isolated topography.*, 2006. "Laurea" Thesis, University of Rome "La Sapienza", Rome, Italy.
- [7] P. B. RHINES, *Geostrophic turbulence*, 11 (1979), pp. 401–441.
- [8] P. SAUNDERS AND B. KING, *Bottom currents derived from a shipborne ADCP on WOCE cruise A11 in the South Atlantic.*, J. Phys. Oceanogr., 25 (1995), pp. 329–347.

Turbulent Mixing in Shear Driven Stratified Fluid

Amrita Shravat

January 12, 2009

1 Introduction

Background turbulence frequently causes mixing across gravitationally stable density interfaces in atmosphere and ocean. Some common examples include turbulence by surface winds or the cooling causing the deepening of the upper ocean mixed layer into stably stratified pycnocline and turbulent mixing at the tropopause causing the growth of a planetary boundary layer. There can be some mixing at the thermocline due to discharge of the buoyant jets from power plants. These mixing processes are important in order to understand and control the biological activities and dispersion of pollutants in the environment ([3]). For example, the diffusion of methane gas in mine shafts.

Many laboratory and associated theoretical studies have considered the response of a stratified fluid to impulsive surface forcing. [3] provides an excellent review on studies with turbulence and vertical shear, investigating the growth of mixed layer in a stratified fluid. The nature of mixing at sheared and shear free density interfaces is very different. For example, the mixing in shear-driven stratified flows, where a well mixed turbulent layer entrains an adjacent non-turbulent layer, is believed to depend upon the external forcing, the depth of the mixed layer and the buoyancy jump across the layer. All these parameters are investigated in the present study of 2-layer fluid system.

The first set of experiments undertaken to explore the growth of the mixed layer as a function of these factors were that of [4]. They found that the entrainment rate was proportional to the friction velocity u_* ($(= |\tau|/\rho_0)^{1/2}$, where τ is the surface shear stress and ρ_0 is the reference density) and Ri_τ^{-1} , where Ri_τ is the overall Richardson number, defined as $g'h/u_*^2$. The current investigation has been motivated by the stratified flow experiments of [2] and [1] in cylindrical geometries. The present work is based on the earlier two-layer model of [1], where a horizontal disk at the base of the tank drives the flow. Their analysis showed that for Richardson Number, $Ri_B (= g'h/\Omega^2 R^2) < 1.5$, the growth of the elevation of the interface separating the upper (almost quiescent) layer of constant density from the lower mixed region is proportional to Ri_B^{-1} . Note that h is the depth of the mixed layer, Ω is the rotation rate of the disk and R is the radius of the disk. The buoyancy is defined as $g' = g(\rho_L - \rho_u)/\bar{\rho}_L$, where g is the gravitational acceleration, ρ_L is the density of the lower layer fluid, ρ_u is the density of the fluid in the upper layer and $\bar{\rho}_L$ is the average density of the fluid in the lower layer, relevant if the fluid in the lower layer is linearly stratified.

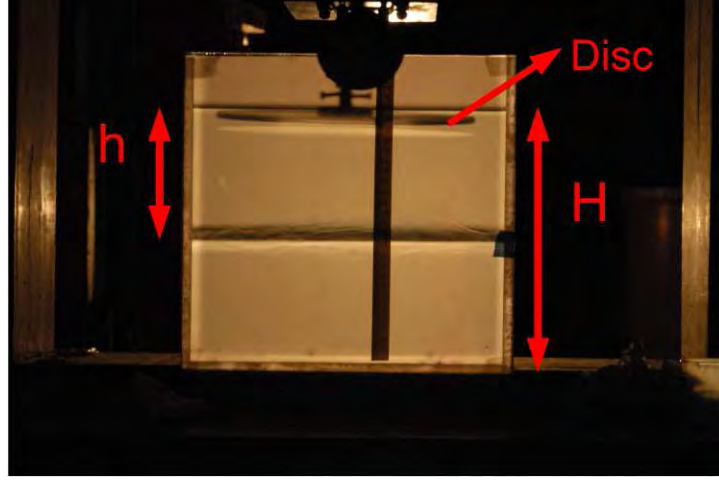


Figure 1: Experimental set-up showing a tank with a horizontal disk spinning the fluid in the upper layer.

The experiments in the current work are characterised by the growth of a mixed layer, separated by the undisturbed stratified lower fluid by a curved surface. In this report, the experimental set-up is discussed, along with the dimensionless parameters that come into play. The experimental results are presented and an attempt to offer a theoretical explanation of these results is made.

2 Experimental set-up and parameters

The experiment comprises of a cylindrical tank of diameter 30cm, with a horizontal disk at the top attached to a motor. Figure (1) and figure (2) show a square tank used initially. The disk is just below the free surface of the top layer, fully immersed in the upper layer fluid. The radius of the disk is 12 cm. The rate of rotation of the disk is controlled by a motor. The rotation rate ranges from $1s^{-1} \leq 2\pi\Omega \leq 12s^{-1}$, where Ω is the angular rotation rate, which remains constant during an experiment. As the disk spins, it provides energy for mixing the two layers. A conductivity probe is used to obtain vertical density profiles in the fluid. The total depth of the fluid, $H = 27\text{cm}$ is fixed and is the same for every experiment. The depth of the upper layer, h , varies during the experiments. The initial depth of the upper layer is indicated as $h_0 = h(t = 0)$, and the depth during an experiment is $h = h(t)$. For most experiments, h_0 is varied between 6cm and 20.6cm. The consequences of changing the Richardson number by varying R , the radius of the disk, are not explored in the present experiments.

The most important dimensionless parameter for the flow is the Richardson number ([1]). The following are the definitions of the Richardson number that will be used throughout this report.

$$Ri_B(t) = \frac{g'h}{\Omega^2 R^2},$$

where R is the radius of the disk, $g' = g(\rho_L - \rho_u)/\bar{\rho}_L$ (ρ_L, ρ_u and $\bar{\rho}_L$ as defined in the previous section), and Ω is the rate of rotation of the disk. Ri_B is constant for most part of the experiment as can be seen from Figure 6. For almost all the experiments, after an initial spin-up (when the fluid in the top layer is set into motion), Ri_B reaches a constant value and maintains that value until the very end of the experiment, where the cylinder bottom affects the experiment. Near the bottom, the boundary layer effects cannot be neglected. Consequently, the value of Ri_B used in this work is \bar{Ri}_B , i.e. the average value of Ri_B calculated ignoring the beginning and the end of an experiment. It should be noted that in this definition, the velocity is assumed to be the “solid body rotation” velocity, i.e. u_u (see below) is given by ΩR . Hence, Ri_B is the same as Ri_τ , discussed in previous section, if the velocity is ΩR .

$$Ri_0 = \frac{g'_0 h_0}{\Omega^2 R^2},$$

where g'_0 is the initial buoyancy jump across the interface between the upper and lower layer, measured before starting the experiment and h_0 is the initial depth of the upper layer.

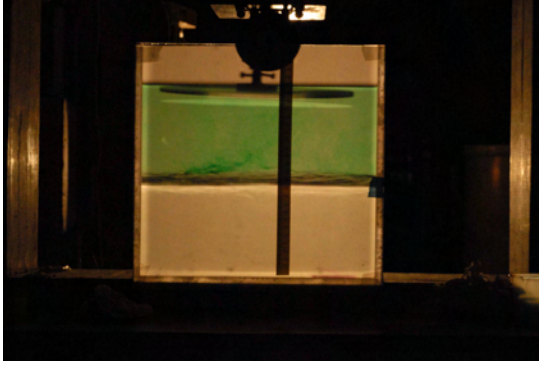
$$Ri_u = \frac{g' h}{u_u^2} = Ri_B \left(\frac{\Omega^2 R^2}{u_u^2} \right), \quad \text{where } u_u \text{ is the characteristic velocity in the upper layer.}$$

$$Ri_m = \frac{g' L_m}{u_u^2},$$

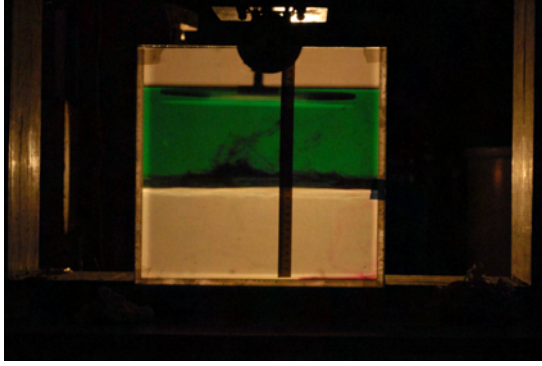
where, L_m is the interfacial length scale (i.e. radius of an eddy at the interface). L_m is proportional to the Ozmidov scale, L_0 . The Reynolds number in the present experiments is sufficiently high for the flow to be regarded as turbulent. Hence, allowing for the assumption that the flow is independent of Reynolds number. The Schmidt number, Sc (ratio of kinematic viscosity to mass transfer diffusion coefficient), is not an important parameter for the flow as $Sc \gg 1$ for a salt stratified water solution and it is constant for all the experiments. Thus, the overall fluid motion and mixing process is assumed to depend on a single parameter, the Richardson number.

2.1 A Typical Experiment

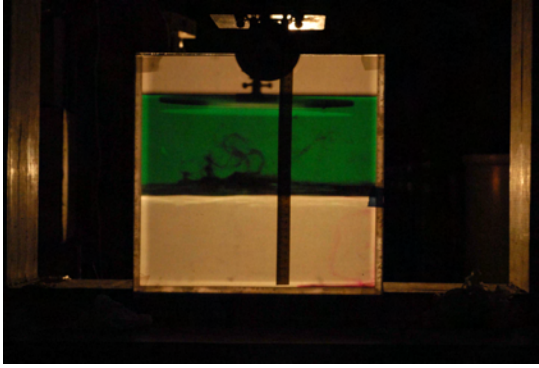
Figure 2 shows a two-layer experiment, with $\bar{Ri}_B = 0.56$. The upper layer is well mixed with fluid being pulled up by the spinning disk, resulting in a dome-shaped interface. The Ekman layer, below the disk, causes the fluid below to rise up. This results in a circulation, where the fluid ascends from the middle and comes down from the side of the tank. As the mixing at the interface occurs, the upper layer deepens until all the fluid in the lower layer is mixed. It is worth noting that the lower layer is very quiescent. It is motionless and retains this property until it is completely mixed. A schematic of the fluid circulation in the upper and lower layer is shown in figure 3. The overturning billows seen at the interface (for the case where $Ri_0 \leq 1$), entrain the lower layer fluid. This mechanism converts the kinetic energy (provided by the disk) of the fluid in the upper layer, into potential energy by lifting up the denser fluid particles from the lower layer and mixing them with the rest of



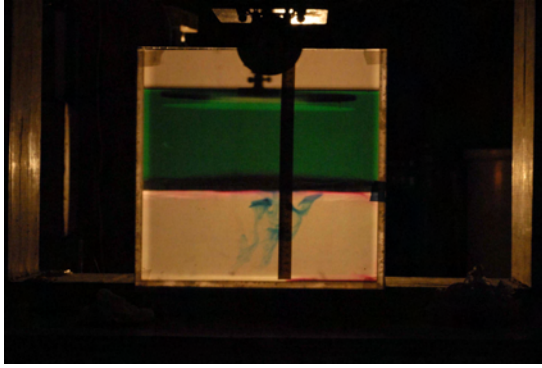
(a) $\hat{t} = 8$



(b) $\hat{t} = 17$



(c) $\hat{t} = 18$



(d) $\hat{t} = 28$

Figure 2: A two-layer experiment with $\Omega = 1s^{-1}$, $\hat{t} = \Omega t$, $Ri_B = 0.56$. In figure 2(c), the fluid in the upper layer is pulled by the disk, spinning at the rate Ω at the top. The fluid in the lower layer is motionless. The dye is rising in 2(d) because it is lighter than the lower layer fluid.

the fluid in the upper layer. It is the size of these overturning billows (in terms of L_m) that determines the rate at which the fluid is mixed. However, once the size exceeds a certain critical value, these billows can no longer develop and the fluid is not mixed efficiently. For $Ri_0 > 1$, a completely different regime is observed. The flow is no longer turbulent and the overturning billows are replaced by waves at the interface.

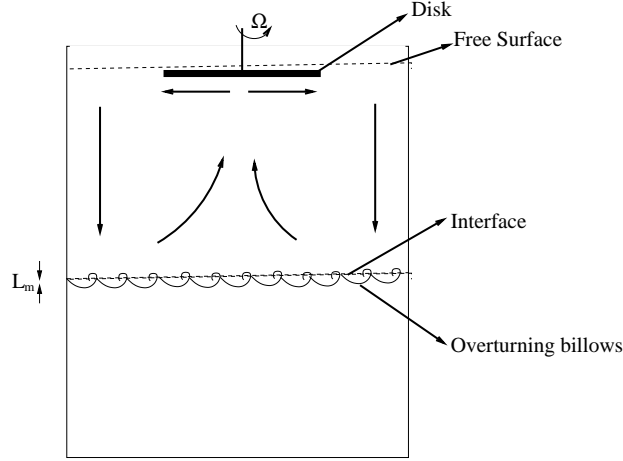


Figure 3: A schematic of the flow in the cylindrical tank, with the disk spinning at the top. There is no circulation in the lower layer fluid.

As mentioned earlier, a conductivity probe is used to obtain vertical profiles of the fluid's density. Figure 4, shows one such profile with $\bar{Ri}_B = 0.19$. The height of the interface (in mm) is plotted against the density of the fluid (in g/cm^3). The evolution in time of the profiles is from left to right. The profile on the extreme left represents the initial density of the two layers. After the disk starts spinning, the first density profile obtained is the second profile from the left (blue line). It is clear, from looking at the profiles from left to right, that the upper layer is deepening as the level of the interface is going down in each subsequent profile. As expected, the lower layer maintains its density until it is completely mixed with the upper layer fluid.

3 Experiments

The following table shows the various experiments conducted to observe the mixing in a two-layer stratified fluid.

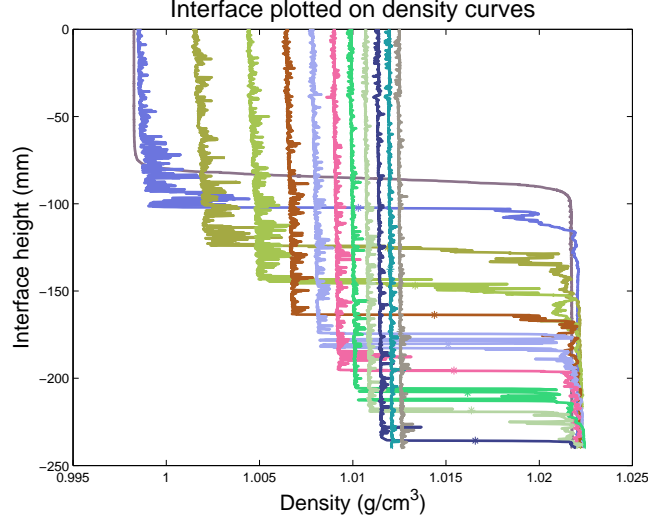


Figure 4: A typical vertical profile of the density during an experiment with $\bar{Ri}_B = 0.19$. The evolution of the profiles in time is from left to right.

Experiment	Ri_0	h_0 (cm)	$\Omega(s^{-1})$	$g'_0(cm s^{-2})$	Ri_B
080713lo40	0.1527	8.7	3	22.7407	0.19
080714lo40	0.3109	17.8	3	22.6470	0.32
080715lo40	0.2367	13.5	3	22.7215	0.255
080716lo40	0.2554	18.5	3	17.8929	0.275
080718lo30	0.2406	6.1	2	22.7218	0.265
080720lo64	0.2336	13.5	5.4	73.1874	0.258
080721lo60	0.0851	13.5	5	22.7028	0.096
080727lo20	2.1293	13.5	1	22.7124	2.1
080731lo30	0.5317	13.5	2	22.6842	0.56
080809lo50	0.1328	13.5	4	22.6561	0.154
080810lo28	1.0016	20.6	1.8	22.6842	0.935
080812lo40	0.2124	13.5	3	20.3892	0.24

The last experiment, ‘080812lo40’, has a buoyant upper layer and a linearly stratified lower layer with density profile evolving in time as shown in figure 5. The range of Richardson numbers explored in the experiments is shown in figure 6.

As expected, mass is conserved in all the experiments. This is clear from figure 7. The value of $Ri_B/Ri_0 = g'h/g'_0h_0$ is approximately 1 for the experiments, implying that $g'h$ is a constant. The fluctuations seen in figure 7 are due to the fact that the flow is turbulent. It is for this reason that the value of Ri_B/Ri_0 is not exactly 1. Another observation that can be made from this figure is that the value of Ri_B/Ri_0 decreases dramatically for some cases but not for the others. This can be attributed to the fact that all the profiles are included in these experiments as there were relatively few of them. Since Ri_0 for these experiments is small, the mixing takes place very quickly as there is a small density difference between the two layers. As mentioned earlier, the value of Ri_B/Ri_0 at the bottom cannot be trusted

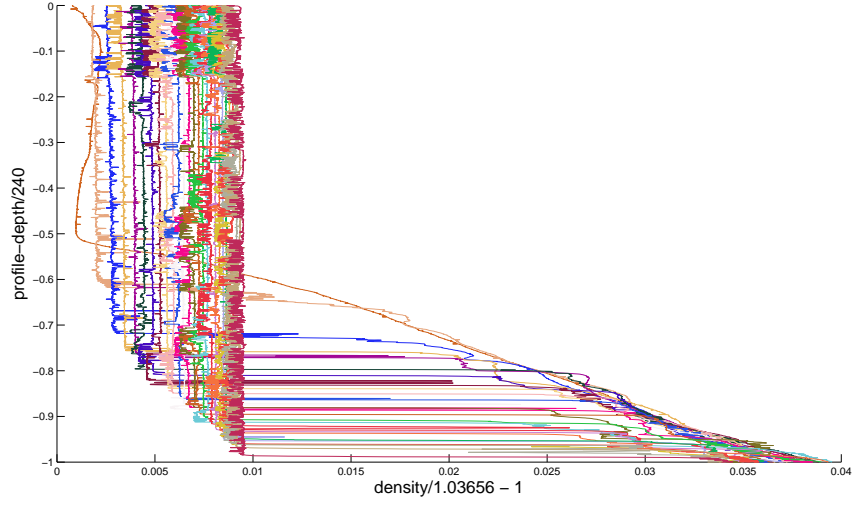


Figure 5: Density Profiles from experiment '080812lo40': Plot of density (non-dimensional) against depth that has been profiled (non-dimensional). Note the profiling depth is not the same as the total depth, $H = 27\text{cm}$, instead the profiling depth is 24cm . Profiles evolve in time from left to right. $\bar{Ri}_B = 0.24$.

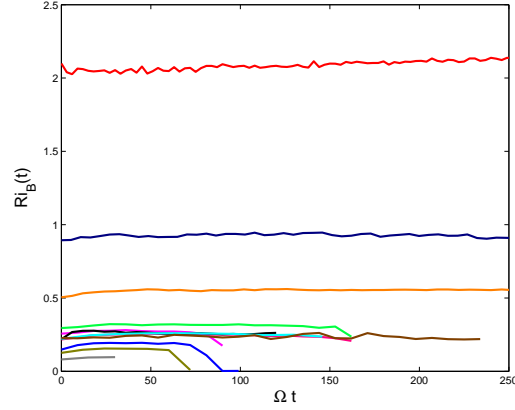


Figure 6: Range of $Ri_B(t)$ explored in the experiments.

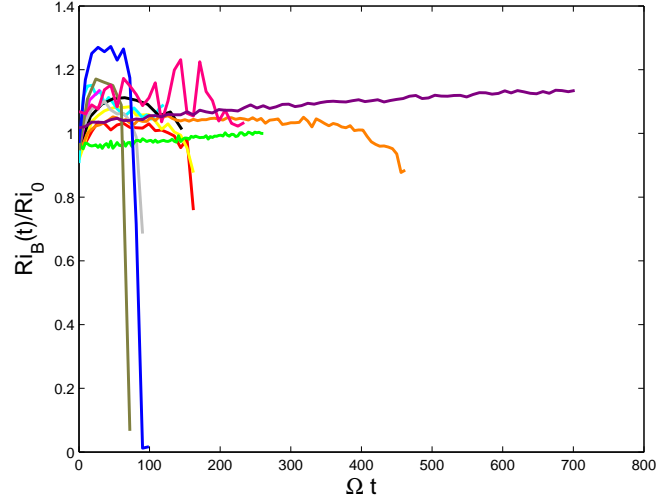


Figure 7: The value of $Ri_B(t)/Ri_0 = g'h/g'_0h_0 \sim 1$. This suggests that $g'h$ is a constant.

because of the effect of bottom boundary layer. Hence, the last few profiles show such a dramatic change in the value of $Ri_B(t)$. Moreover, the initial value for almost all the experiments is 1. It changes after the initial spin-up of the fluid, again reinforcing that after the initial spin-up period, the value of Ri_B is almost a constant. From now onwards, it is this value, found by averaging the value of $Ri_B(t)$ (excluding the beginning and the end profiles) that will be used for Ri_B .

It can be established from the following theoretical consideration that $g'h$ is a constant. If A is the cross-sectional area of the cylinder, H is the total height, h is the height of the upper layer, which evolves in time $h = h(t)$, and Q = entrainment flux, then

$$\frac{d}{dt}(Ah) = Q. \quad (1)$$

The amount of fluid entrained is

$$\frac{d}{dt} \left(A \int_0^h \rho_u dz \right) = Q\rho_L, \quad (2)$$

where ρ_L is the density in the lower layer. The average density in the upper layer is

$$\bar{\rho}_u = \frac{1}{h} \int_0^h \rho_u dz, \quad (3)$$

Hence, on substituting (3) into (2),

$$\frac{d}{dt}(Ah\bar{\rho}_u) = Q\rho_L. \quad (4)$$

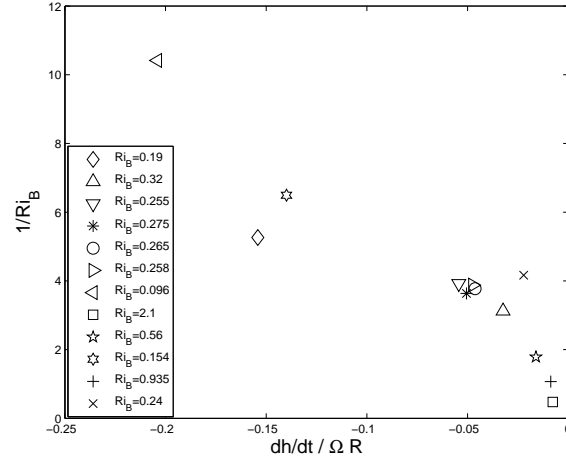


Figure 8: Plot of $dh/dt/\Omega R$ against $1/Ri_B$. The growth of the interface is not perfectly proportional to $1/Ri_B$.

which gives

$$\frac{d}{dt}[A(\bar{\rho}_u - \rho_L)h] + \rho_L \frac{d}{dt}[Ah] = Q\rho_L, \quad (5)$$

$$\frac{d}{dt}[A(\bar{\rho}_u - \rho_L)h] = 0, \quad (6)$$

$$\frac{d}{dt}\left[Ag \frac{(\bar{\rho}_u - \rho_L)}{\rho_L} h\right] = 0, \quad (7)$$

This shows that $\frac{d}{dt}(g'h) = 0$ and thus $g'h$ is a constant and mass is conserved.

In figure 8, the growth of the interface (scaled by ΩR) is plotted against $1/Ri_B$. This figure suggests that the mixing, or the growth of the interface, does not scale perfectly with Ri_B^{-1} . From the experimental results presented here, it is clear that the outcome of this work does not support the conclusion of [1], where they found that the growth of the interface is proportional to Ri_B^{-1} , where the velocity ($= \Omega R$) of the upper layer fluid is constant in each experiment. Thus the simple mixing law, inferred by [1] needs a modification. This indicates that the results do not support the assumption that the fluid is in solid body rotation, i.e. the velocity does not scale as ΩR as assumed in the definition of Ri_B . Hence, it is not completely accurate to assume that the velocity in the upper layer is the same as ΩR and that it does not change during the experiment. An immediate consequence of this result will be to re-examine the velocity of the upper layer fluid.

4 Theory

The driving disk sets up a characteristic velocity u_u in the upper layer. Assuming that the bottom layer is stationary, the interfacial stress, τ can be described by the following expression,

$$\tau = c_D \rho_L u_u^2,$$

where c_D is the drag coefficient. Hence, the work done, W , can be expressed as

$$\frac{dW}{dt} = \pi R^2 u_u \tau = \pi R^2 c_D \rho_L u_u^3, \quad (8)$$

The potential energy, PE is

$$PE = \pi R^2 \left[-g \int_0^h \rho_u z dz - g \int_h^H \rho_L z dz \right] = (\pi R^2 \rho_L) \frac{g' h^2}{2} + k, \quad (9)$$

It should be noted here that k is a constant. However, the reference level for potential energy can always be chosen such that this constant k is zero. The potential energy is negative because $z = 0$ is defined at the top of the tank. The rate of change of potential energy of the fluid parcels is proportional to the amount of power provided by the disk, i.e.

$$\frac{d}{dt} PE = \Gamma \frac{dW}{dt}, \quad (10)$$

where Γ is the flux coefficient ([5]). Assuming that $g'h$ is a constant (see previous section), this leads to

$$\frac{d}{dt} \left[(g'h) (\pi R^2 \rho_L) \frac{h}{2} \right] = \Gamma \pi R^2 c_D \rho_L u_u^3. \quad (11)$$

Hence, the interface grows as

$$\frac{dh}{dt} = 2\Gamma \frac{c_D u_u^3}{g'h}. \quad (12)$$

Let

$$\hat{h} = \frac{h}{h_0}, \quad \hat{t} = \Omega t, \quad (13)$$

$$\frac{d\hat{h}}{d\hat{t}} = \frac{2c_D \Gamma u_u^3}{\Omega h_0 g'h}.$$

It is quite intuitive to expect that the growth of the interface depends on the amount of energy that is spent on mixing the two layers together. The following 3 sub-sections quantify this statement based on the amount of energy (provided by the disk) that is spent on entrainment.

4.1 Case 1

Solid Body Rotation:

If none of the energy is spent on entrainment, then all of the energy provided by the disk is exhausted in moving the fluid in the upper layer. The upper layer in this case moves with a constant speed, ΩR , provided by the disk. In this case, the upper layer fluid is in solid body rotation, i.e.

$$u_u = \Omega R,$$

where u_u is the speed of the upper layer fluid.

$$\frac{d\hat{h}}{d\hat{t}} = \frac{R}{h_0} \frac{c_1}{Ri_B} = \text{constant}, \quad (14)$$

$$c_1 = 2c_D\Gamma, \quad (15)$$

$$\hat{h} - 1 = \frac{R}{h_0} \frac{c_1}{Ri_B} \hat{t}. \quad (16)$$

This case is similar to the outcome of [1], with $\frac{d\hat{h}}{d\hat{t}}$ a constant.

4.2 Case 2

ALL Energy spent on entrainment:

Assume u_u can be expressed as

$$u_u = \Omega R \left(\frac{h_0}{h} \right)^\alpha,$$

where h_0 is the initial depth of the upper layer measured before starting the experiment. This expression means that the velocity of the upper layer fluid decreases with time. It is no longer a constant, which is the case if some energy is spent on mixing the two layers. The total kinetic energy, KE_u in the upper layer is

$$KE_u = \frac{\pi R^2}{2} \rho_{ref} u_u^2 h, \quad (17)$$

where ρ_{ref} is the density taken at some reference level. This reference density can, for example, be the average upper layer density. It is assumed that in this case all the energy is spent on mixing. This signifies that the kinetic energy of the upper layer does not change as all the energy that the disk provides is lost in mixing the two layers. Hence,

$$\frac{d}{dt} KE_u = 0,$$

Thus,

$$\frac{\pi R^2}{2} \rho_{ref} u_u^2 h = \text{constant}, \quad (18)$$

By appealing to the Boussinesq approximation, changes in ρ_{ref} can be ignored. Thus, $u_u^2 h$ is constant if $\alpha = 0.5$ i.e

$$u_u = \Omega R \left(\frac{h_0}{h} \right)^{1/2}.$$

Hence, for the above expression to hold true, the fluid must spend all the kinetic energy on mixing.

4.3 Case 3

Intermediate Case:

The two cases mentioned above are the extreme cases when either all the energy from the disk is spent on mixing or all of it is expended on spinning the fluid in the upper layer. A more sensible scenario is if part of the energy from the disk is consumed by the mixing process and part of it is spent on mobilising the fluid. Suppose, as before,

$$u_u = \Omega R \left(\frac{h_0}{h} \right)^\alpha.$$

Substituting this expression into equation (13) and integrating yields the following expression for \hat{h} :

$$\hat{h} - 1 = \left[\frac{R}{h_0} \frac{(1 + 3\alpha)c_1}{Ri_B} \hat{t} + 1 \right]^{1/(1+3\alpha)} - 1. \quad (19)$$

Thus we can summarise the above results as follows:

- If $\alpha = 0$: Entrainment has no effect on energy - this is the case when stratification is weak.
- If $\alpha = 0.5$: All the kinetic energy from the disk is spent on entraining the lower layer fluid. Hence, there can be three dynamical regimes based on the value of Ri_B . When Ri_B is small, it can be expected that the fluid is more or less in solid body rotation. When the Richardson number is small, it is relatively easy for the upper fluid parcels to lift the lower fluid and mix. This means that more energy is left for mobilising the upper layer fluid. When Ri_B is large, it is presumed that most of the energy would be spent on lifting up the lower layer fluid. For moderate values of Ri_B , some of the energy is spent on the potential energy of the lower layer fluid parcels and a part is spent on mobilising fluid in the upper layer. This is easier to visualise from figure 9. In the three regimes shown, the velocity at the beginning is ΩR . Ri_B increases from left to right in the figure, where (a) shows that the velocity of the fluid in the upper layer is a constant as for small Ri_B most of the energy is spent on mobilising the upper layer fluid. The middle picture, (b) represents the intermediate case where the velocity is not a constant and the third case, (c) represents the large Ri_B regime. The flow in this regime is not as turbulent as the other cases. There are no overturning billows at the interface, instead it has waves. In this case, almost all the energy is exhausted in mixing the fluid layers rather than providing kinetic energy to the upper

layer fluid. The rate of change of interface drops from left to right. As expected, when stratification is weak, it is relatively easy to mix the two fluids together. However, as the stratification increases, it becomes increasingly expensive to mix the fluid. In the high Richardson number case, the mixing is very inefficient, because of the lack of overturning billows. Hence, it takes a significantly long time for the two layers to mix.

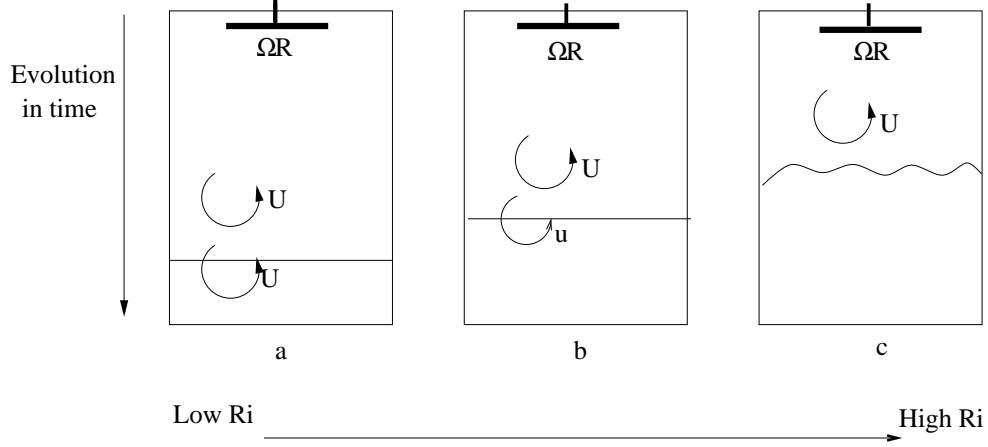
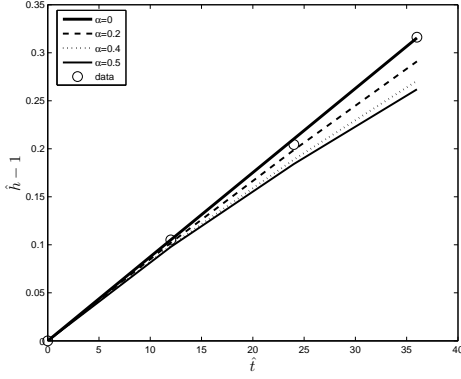


Figure 9: A sketch of the three dynamical regimes based on the Richardson number (see text). The Richardson number increases from left to right in the figure. In b, $u < U < \Omega R$ as part of the energy is spent on mixing the fluid and only part of it is used to mobilise the upper layer fluid.

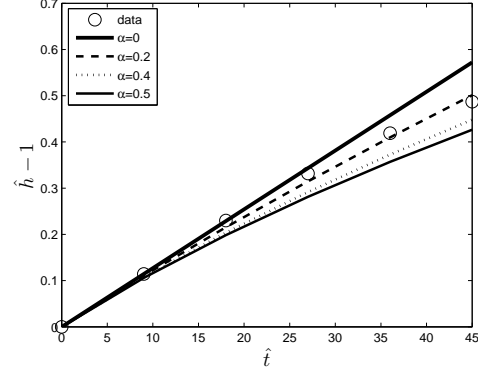
5 Discussion

In order to analyse the data to test the three regimes scenario that have been discussed qualitatively, equation (19) is plotted for various values of α , with $\hat{h} - 1$ against \hat{t} . The constant c_1 is calculated from the initial data gathered just after the spin-up stage. It is assumed that just after the spin-up, the upper layer fluid is in solid body rotation. This is a reasonable assumption because at the beginning the work is done mainly on mobilising the upper layer fluid. The initial data (excluding the spin-up stage) is used to plot equation (16). A value of c_1 can be found from the slope of this curve. This constant c_1 is assumed constant throughout the experiment. The value of c_1 is employed to plot $\alpha = 0.2, 0.4$ and 0.5 curves. The data in figure 10 is very close to the line $\alpha = 0$, as expected from the energy arguments. The stratification in these experiments is quite weak, which suggests that the mixing requires very little energy and most of the kinetic energy from the disk is consumed by the fluid in the upper layer.

Similarly, for the curves in figure 11, the data lies above the line $\alpha = 0.5$, indicative of an intermediate regime. Figure 11(e), shows a good agreement with the theoretical arguments. The data lies just above the $\alpha = 0.5$ curve, again indicating that the upper layer fluid is



(a) $Ri_B = 0.15, c_1 = -0.0020$



(b) $Ri_B = 0.19, c_1 = -0.0030$

Figure 10: Equation (19) plotted with $\hat{h} - 1$ against \hat{t} . It is plotted for $\alpha = 0, 0.2, 0.4, 0.5$.

not in solid body rotation and the velocity decreases in time. For this moderate value of Richardson number, relatively more energy is required to mix the two layers. As a result, lesser energy is spent on driving the upper layer fluid. In figure 11(f), the data is very close to $\alpha = 0.5$ curve, indicating that most of the energy is consumed by the process of mixing.

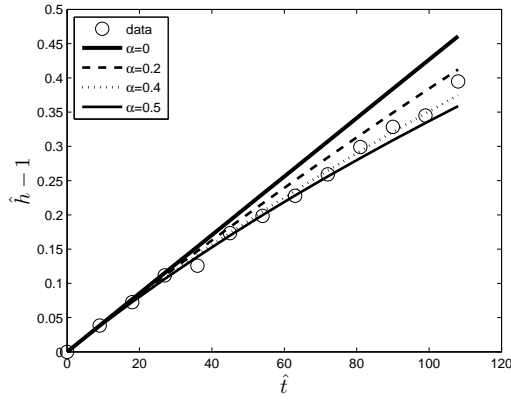
Figure 12 shows the plot of $\hat{h} - 1$ against \hat{t} from equation (19) for an experiment with two-layers, where the lower layer is linearly stratified. Figure 5 shows its vertical density profiles. It appears from figure 12 that the data does not fit the theoretical model. It should be noted, however, that the constant c_1 needs to be recalculated after every profile. This has not been done yet. Qualitatively, it can be seen that at the beginning, when g' is very small, the data is close to the $\alpha = 0$ curve. As g' increases, it gets closer to the $\alpha = 0.5$ line, as expected. This is consistent with the argument that the larger the value of Ri_B or stronger the stratification, the harder it is for the upper layer fluid to mix the denser lower layer. Consequently, most of the energy is spent on mixing the fluids.

The flux coefficient, Γ , (defined in section 4) can be divided into 3 components: Γ_u : mobilising flux coefficient, Γ_D : dissipation flux coefficient, Γ_ρ : Potential Energy flux coefficient. Thus, Γ_u determines the amount of energy from the disk going into mobilising the fluid in the upper layer, Γ_D determines the energy spent on dissipation and Γ_ρ , the energy that is used up in lifting the fluid parcels from below. The reason for dividing Γ into three components is to understand and establish the fraction of energy extracted from the disk to drive the upper layer fluid, mix the two layers together and dissipate.

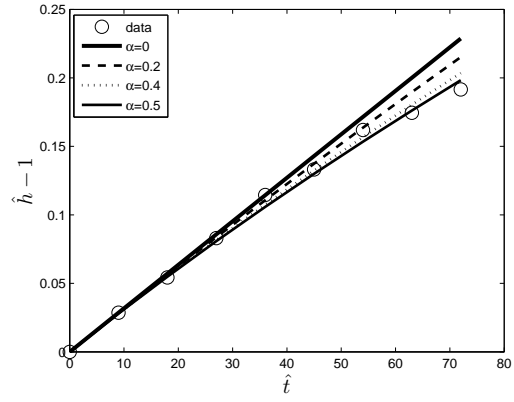
The growth of the interface occurs due to entraining of the lower layer fluid. Thus, equation (12) can be written more precisely in terms of Γ_ρ as

$$\frac{dh}{dt} = 2\Gamma_\rho c_D \frac{u_u^3}{g'h}, \quad (20)$$

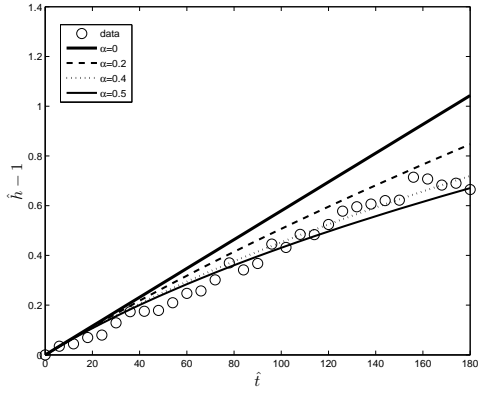
$$\frac{dh}{dt} = \frac{2\Gamma_\rho c_D}{Ri_m} \frac{L_m}{h} u_u, \quad L_m = Ri_m \frac{u_u^2}{g'}, \quad (21)$$



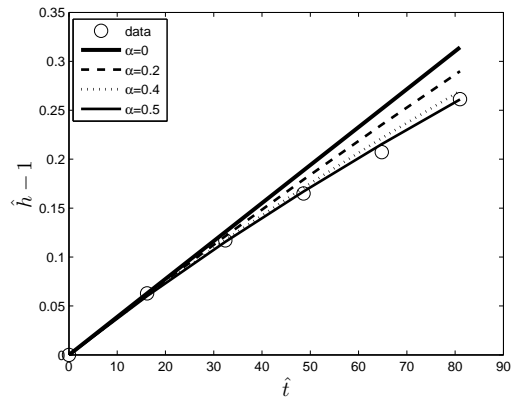
(a) $Ri_B = 0.255, c_1 = -0.0015$



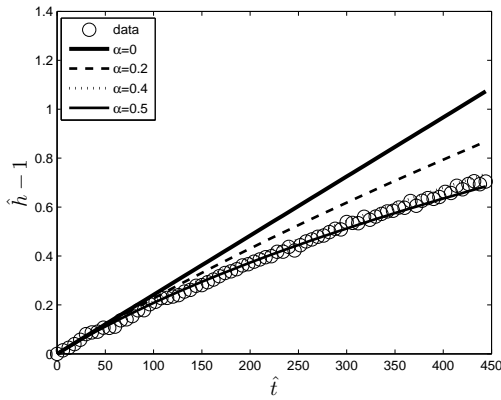
(b) $Ri_B = 0.275, c_1 = -0.0015$



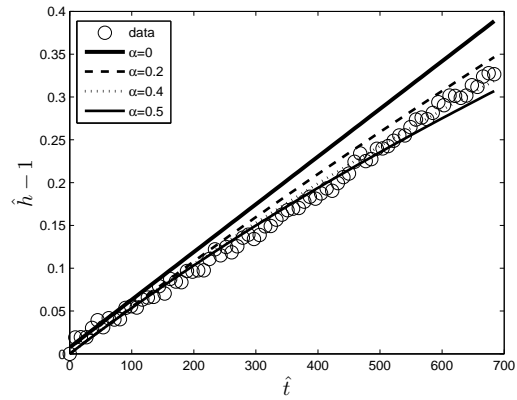
(c) $Ri_B = 0.265, c_1 = -0.0017$



(d) $Ri_B = 0.258, c_1 = -0.0013$



(e) $Ri_B = 0.56, c_1 = -0.0015$



(f) $Ri_B = 0.935, c_1 = -0.00086$

Figure 11: Plot of $\hat{h} - 1$ against \hat{t} with value of α ranging from 0 to 0.5. The data lies above $\alpha = 0.5$ curve, suggesting that these experiments fall under the intermediate regime.

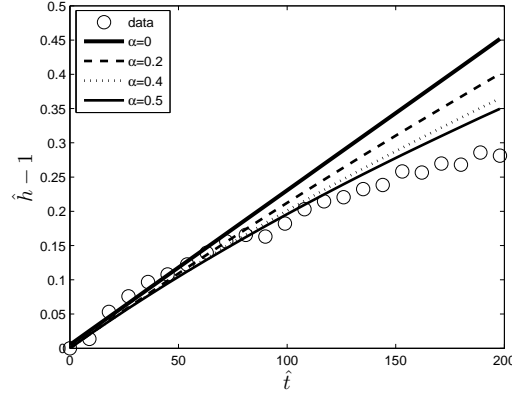


Figure 12: Linear Stratification, $Ri = 0.24$, $c_1 = -0.00082$

Remember L_m is the interfacial length scale. From the definition of c_1 , given by (15), it seems reasonable that it should be defined more precisely as $c_1 = 2c_D\Gamma_u$, since it is the mobilising component of flux that plays an important role here. As discussed before, the kinetic energy provided by the rotating disk is spent on dissipation, increasing the kinetic energy of the upper layer by mobilising it and increasing the potential energy of the flow by lifting the heavier lower layer particles and mixing them with the upper layer. Since the flow has high Re ,

$$\epsilon = \frac{a_2 u_u^3}{h}$$

(from classical turbulence theory)

Parametrizing the change in kinetic energy of the two-layer fluid by ϵ , produces the following equation

$$\frac{d}{dt} \left(\frac{u_u^2 h}{2} \right) = a_1 \epsilon L_m (= -F - D + M), \quad (22)$$

where F is the force responsible for changing the potential energy of the flow, D is the dissipation and M is the force applied on the upper layer for driving the fluid. a_1 is a scaling parameter.

$$\frac{d}{dt} \left(\frac{u_u^2 h}{2} \right) = 2\Gamma_u c_D \frac{L_m}{h} \frac{u_u^2}{2} u_u, \quad (23)$$

if $a_1 a_2 = \Gamma_u c_D$. From (21),

$$c_D \frac{L_m}{h} u_u = \frac{Ri_m}{2\Gamma_\rho} \frac{dh}{dt}, \quad (24)$$

$$\frac{d}{dt} \left(\frac{u_u^2 h}{2} \right) = \frac{u_u^2}{2} \frac{2\Gamma_u Ri_m}{2\Gamma_\rho} \frac{dh}{dt}. \quad (25)$$

Assume,

$$\frac{u_u^2}{2} = \frac{\Omega^2 R^2}{2} \left(\frac{h_0}{h} \right)^{2\alpha},$$

$$\left(1 - \frac{2Ri_m \Gamma_u}{2\Gamma_\rho} \right) \frac{u_u^2}{2} \frac{dh}{dt} - 2\alpha \frac{\Omega^2 R^2}{2} h_0^{2\alpha} h h^{-2\alpha-1} \frac{dh}{dt} = 0, \quad (26)$$

$$\left(1 - \frac{2Ri_m \Gamma_u}{2\Gamma_\rho} \right) \frac{u_u^2}{2} \frac{dh}{dt} - 2\alpha \frac{u_u^2}{2} \frac{dh}{dt} = 0, \quad (27)$$

$$\frac{Ri_m \Gamma_u}{\Gamma_\rho} = 1 - 2\alpha. \quad (28)$$

It has been found experimentally that $Ri_m = 0.2$ (see [3], [6]). Substituting $\alpha = 0.4$ (as suggested by figure 11(e)) along with $Ri_m = 0.2$, analysis reveals that

$$\Gamma_u \sim \Gamma_\rho.$$

This is an interesting result showing that for intermediate values of Ri_B , an equal amount of energy goes into driving the fluid and in mixing the layers.

Moreover, if $\alpha = 0, \Gamma_u = 5\Gamma_\rho$. Thus, for the solid body rotation case (small Ri_B), the proportion of energy that goes into driving the fluid is much larger than the proportion that is spent on mixing. On the other hand, if $\alpha = 0.5, \Gamma_u = 0$. As mentioned before, this represents the case where there is no change in the kinetic energy of the upper layer fluid. Thus, all the energy is spent on entraining the lower layer fluid.

6 Conclusion

It can be concluded from the above analysis and experimental work that a Ri_B^{-1} law for mixing is not found. Moreover, the fluid velocity in the upper layer is not always ΩR . It would appear that the upper layer is not always spun-up during the experiment and is not constant with h . It can be seen both from the theory and the density profiles that the height of the interface decreases with time. For moderate values of Richardson number, energy spent on mixing is of the same order as the energy that goes into mean velocity. For values of $Ri_B \geq 1$, the dynamics is completely different. The flow is no longer turbulent. This implies that the Reynolds number can play a significant role in determining the dynamics and can no longer be ignored. The qualitative behavior observed during an experiment at these values of Ri_B is very different than the observed behavior for lower values of Ri_B . The fluid is no longer mixed by eddies at the interface. The interface in this case is very thin and dominated by wave activity.

It is safe to say that the model developed here can capture the various aspects of mixing for $Ri_B \leq 1$. More work is required to understand the higher Richardson number cases.

7 Acknowledgements

This work was funded under the Geophysical Fluid Dynamics Summer Program at Woods Hole Oceanographic Institution. A special thanks to Colm Caulfield and Claudia Cenedese for the inspiration and guidance of this project. The many discussions with them were extremely valuable and helpful. Thanks to Neil Balmforth for directing the summer school and all the faculty members who came and made this summer school a very exciting place to learn about the many aspects of geophysical fluid dynamics.

References

- [1] D. BOYER, P. DAVIES, AND Y. GUO, *Mixing of a two-layer stratified fluid by a rotating disc.*, Fluid Dynamics Research, 21 (1997), pp. 381–401.
- [2] P. DAVIES, Y. GUO, D. BOYER, AND A. M. FOLKARD, *The flow generated by the rotation of a horizontal disk in a stratified fluid.*, Fluid Dynamics Research, 17 (1995), pp. 27–47.
- [3] H. FERNANDO, *Turbulent mixing in stratified fluids.*, Annu. Rev. Fluid Mech., 23 (1991), pp. 455–493.
- [4] H. KATO AND O. PHILLIPS, *On turbulent entrainment at a stable density interface.*, J. Fluid Mech., 37 (1969), pp. 643–655.
- [5] T. OSBORN, *Estimates of the local rate of vertical diffusion from dissipation measurements*, Journal of Physical Oceanography, 10 (1980), p. 8389.
- [6] W. PELTIER AND C. CAULFIELD, *Mixing efficiency in stratified shear flows.*, Ann. Rev. Fluid Mech, 35 (2003), pp. 135–124.

Crumpling of a Thin Ice Sheet due to Incident Flow

Toby Wood

Woods Hole Geophysical Fluid Dynamics Program 2008

January 12, 2009

1 Introduction

Consider a sea partially covered by a thin ice sheet, with a prevailing current driving water under the ice edge. If the ice sheet is fixed relative to the flow then a Blasius boundary layer will develop beneath, due to the no-slip boundary condition at the underside of the sheet. This viscous boundary layer will divert the oncoming flow, causing it to run deeper under the ice, and thus produce a region of high pressure within the fluid in the vicinity of the ice edge. This situation is illustrated in Figure 1. The pressure will cause

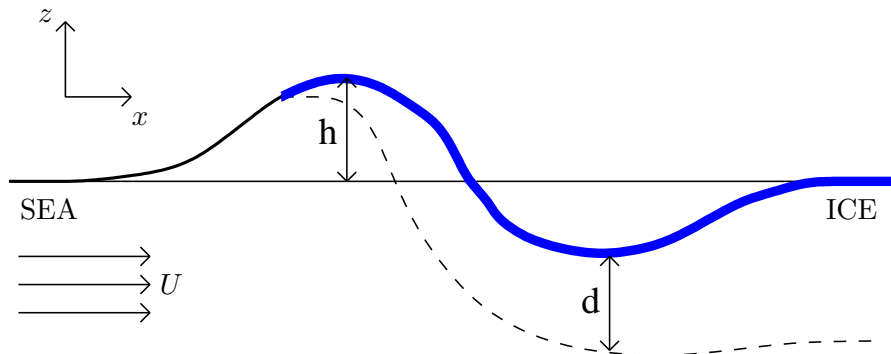


Figure 1: Flow beneath an ice sheet leads to a Blasius boundary layer. The bottom of the boundary layer is indicated by a dashed line.

deformations of both the sea surface and the ice sheet close to the ice edge. If the flow is sufficiently rapid, and the sheet sufficiently thin, then the ice will fracture. Even if the sheet remains intact, it may crumple when the flow speed exceeds some critical value. This crumpling results from a standing flexural wave propagating downstream from the ice edge.

In this report we seek to describe mathematically the deformation of the ice sheet arising from a steady incident flow at the ice edge. (The problem is equivalent to an ice sheet moving across the surface of a stationary fluid.) The formulation of the problem is closely analogous to that employed by Harper & Dixon [3] in their description of the so-called

“Reynolds ridge”. In that problem a strong horizontal gradient in surface tension, due to the presence of surfactant contamination, produces an effective no-slip boundary condition beneath the contaminated part of the free surface. Again, a Blasius boundary layer forms, leading to deformation of the free surface. In the present problem, the bending stiffness (or *flexural rigidity*) of the ice sheet replaces surface tension in limiting the curvature of the surface.

There is also an important analogy between the present problem and that of wave scattering by the ice edge, as studied by Balmforth & Craster [1]. In that work, the “forcing” applied to the ice sheet arises from a wave on the sea surface that is incident on the ice edge. No background flow was included, and so there was no Blasius boundary layer — in fact the fluid was regarded as perfectly inviscid. The inclusion of a background flow in the present problem introduces advection terms into the linearised boundary conditions, and so standing waves appear in the solution when the background flow exceeds a critical value.

2 Formulation

Following Harper & Dixon, we regard the Blasius boundary layer beneath the ice sheet as akin to a thin aerofoil; that is, the bottom of the boundary layer represents an impenetrable barrier to the incompressible fluid flow beneath. Outside the boundary layer we neglect viscous effects, so that the flow may be regarded as irrotational. We may therefore describe the fluid velocity \mathbf{u} in terms of a velocity potential ϕ that satisfies

$$\nabla^2 \phi = 0, \quad (1)$$

$$\mathbf{u} = \nabla \phi. \quad (2)$$

The flow upstream is at speed U in the x direction and z is the vertical coordinate relative to the undisturbed free surface. With regard to Figure 1, we take $h(x, t)$ to be the height of the free surface, and $d(x)$ to be the thickness of the viscous boundary layer. If the ice edge is located at $x = 0$ then

$$d(x) = \begin{cases} 0, & \text{for } x < 0 \\ 1.7208 (\nu x / U)^{1/2}, & \text{for } x > 0 \end{cases} \quad (3)$$

where ν is the kinematic viscosity. Provided that the deflection of the free surface remains small, we may linearise our equations about a state of uniform flow beneath a horizontal free surface. We also suppose that the viscous boundary layer follows the contours of the free surface; this assumption will cease to be valid if the curvature of the free surface becomes comparable to the thickness of the boundary layer [2]. The linearised boundary conditions on the ice surface ($x > 0$) are

$$(\partial_t + U \partial_x)(h - d) = \partial_z \phi|_{z=0} \quad (4)$$

$$(\partial_t + U \partial_x) \phi|_{z=0} + gh = -\frac{B}{\rho} \partial_x^4 h, \quad (5)$$

where ρ is the fluid density and B is the bending stiffness of the ice [5], which is related to the Young's modulus E and Poisson ratio r by

$$B = \frac{E\Delta^3}{12(1-r^2)}. \quad (6)$$

Here, Δ is the thickness of the ice, which we shall assume to be $\approx 1\text{m}$. The characteristic horizontal scale for deformations of the ice surface is the bending length, $L \equiv (B/\rho g)^{1/4}$; for ice of thickness 1m, the bending length is approximately 50m [1].

On the free surface of the open sea, the bending stress vanishes, and in its place we substitute surface tension, σ . We therefore replace (5) with a new dynamic boundary condition for $x < 0$:

$$(\partial_t + U\partial_x)\phi|_{z=0} + gh = +\frac{\sigma}{\rho}\partial_x^2 h. \quad (7)$$

The characteristic horizontal scale on the sea surface is given by the capillary length, $l \equiv (\sigma/\rho g)^{1/2} \approx 3\text{mm}$. Since $l \ll L$ we do not expect surface tension effects to be significant in determining the shape of the ice surface. We have chosen to retain surface tension firstly to regularise any small scale effects occurring on the sea surface. We also have in mind possible laboratory experiments using materials with a much smaller bending length than sea ice, for which surface tension effects would not be negligible.

Due to the discontinuity at the ice edge, we shall need to specify additional ‘‘edge conditions’’ between $x = 0_-$ and $x = 0_+$. In particular, we balance forces and torques at the ice edge by integrating (5, 7) and their first moments across an interval $(-\epsilon, +\epsilon)$, then allowing ϵ to tend to zero. Assuming that \mathbf{u} and h are continuous, we obtain the edge conditions

$$B\partial_x^3 h|_{0_+} = -\sigma\partial_x h|_{0_-} \quad (8)$$

$$B\partial_x^2 h|_{0_+} = -\sigma h|_{0_-}. \quad (9)$$

We now seek two dimensional, steady state solutions of the governing equation (1) subject to (4, 8, 9) and the two-part boundary condition

$$U\partial_x\phi|_{z=0} + gh = \begin{cases} +gl^2\partial_x^2 h, & \text{for } x < 0 \\ -gL^4\partial_x^4 h, & \text{for } x > 0 \end{cases} \quad (10)$$

Following Balmforth & Craster [1], we shall apply the Wiener–Hopf technique, and so we define half-range transforms

$$\Phi_+(k, z) = \int_{0_+}^{+\infty} \phi(x, z)e^{ikx} dx \quad (11)$$

$$\Phi_-(k, z) = \int_{-\infty}^{0_-} \phi(x, z)e^{ikx} dx \quad (12)$$

The full Fourier transform is then $\Phi = \Phi_+ + \Phi_-$ (similar notation will be employed for the transforms of d and h). With these conventions Φ_+ is well defined (and analytic) in the upper half of the complex k -plane, and Φ_- is similarly well defined in the lower half of

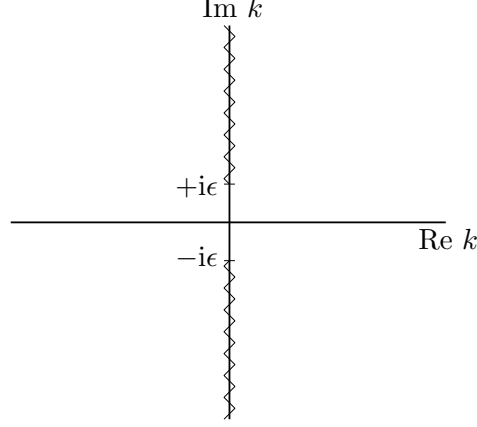


Figure 2: Branch cuts in the complex k plane used in the definition of $\Gamma(k)$.

the complex k -plane. We suppose that Φ_+ and Φ_- can be analytically continued over the complex domains \oplus and \ominus , which include, respectively, the upper and lower half k -planes. The full Fourier transform of ϕ (and of h , d , etc.) is then well defined in the intersection¹ of \oplus and \ominus . We now transform each of our equations in turn. From (1) we find

$$\partial_z^2 \Phi_{\pm} = k^2 \Phi_{\pm} \pm (\partial_x - ik)\phi|_{0\pm} \quad (13)$$

$$\Rightarrow \partial_z^2 \Phi = k^2 \Phi \quad (14)$$

For simplicity, we restrict attention to the case of deep water. We then have a “lower boundary” condition that $\phi \rightarrow 0$ as $z \rightarrow -\infty$. The solutions of (14) may then be written as

$$\Phi(k, z) = \hat{\Phi}(k) e^{\Gamma(k)z}, \quad (15)$$

where $\Gamma(k)$ is defined so that $\Gamma(k) = \pm k$ and $\text{Re } \Gamma(k) \geq 0$. In what follows we shall require a more formal definition of $\Gamma(k)$. We therefore temporarily allow d (and hence ϕ and h) to have some slow sinusoidal modulation in the y direction, with small wavenumber ϵ . Then equation (14) is replaced by

$$\partial_z^2 \Phi = (k^2 + \epsilon^2) \Phi \quad (16)$$

and so $\Gamma(k) = (k^2 + \epsilon^2)^{1/2}$, where we take the branch cuts to lie in the intervals $[\pm i\epsilon, \pm i\infty)$ (see Figure 2). The two dimensional problem represents the limiting case in which $\epsilon \rightarrow 0$, so we define

$$\Gamma(k) \equiv \lim_{\epsilon \rightarrow 0} (k^2 + \epsilon^2)^{1/2}. \quad (17)$$

We now transform the kinematic boundary condition (4) to find

$$-\partial_z \Phi_{\pm} = ikU(H_{\pm} - D_{\pm}) \pm U(h - d)|_{0\pm} \quad (18)$$

$$\Rightarrow -\Gamma \hat{\Phi} = ikU(H - D) \quad (19)$$

¹Formally, the Wiener–Hopf technique requires the intersection of \oplus and \ominus to be some strip containing the real axis. This and other technical details are not addressed here, but are described in detail in [1].

where we have assumed that the boundary layer thickness d is continuous² at $x = 0$. Lastly, transforming the two-part dynamic boundary condition (10) yields

$$g(1 + L^4 k^4)H_+ = ikU\hat{\Phi}_+ + gP_{\text{ice}}(k) + U\phi|_{0+} \quad (20)$$

$$g(1 + l^2 k^2)H_- = ikU\hat{\Phi}_- + gP_{\text{sea}}(k) - U\phi|_{0-} \quad (21)$$

$$\therefore gH + L^4 k^4 H_+ + l^2 k^2 H_- = ikU\hat{\Phi} + gP(k), \quad (22)$$

where

$$P_{\text{ice}}(k) = L^4(\partial_x^3 - ik\partial_x^2 - k^2\partial_x + ik^3)h|_{0+} \quad (23)$$

$$P_{\text{sea}}(k) = l^2(\partial_x - ik)h|_{0-} \quad (24)$$

$$P(k) = P_{\text{ice}}(k) + P_{\text{sea}}(k) \quad (25)$$

are polynomials in k whose coefficients are the (unknown) values of h and its derivatives at the ice edge. Applying the edge conditions (8–9) we find that

$$P(k) = L^4(-k^2\partial_x + ik^3)h|_{0+}, \quad (26)$$

so P is $O(k^2)$ at $k = 0$. We may now eliminate $\hat{\Phi}$ between equations (19) and (22) to obtain a single equation for the free surface:

$$H + l^2 k^2 H_- + L^4 k^4 H_+ = \frac{U^2}{g} \frac{k^2}{\Gamma} (H - D) + P. \quad (27)$$

In §3 we shall apply the Wiener–Hopf technique to this equation in order to find $H(k)$ and hence obtain the free surface $h(x)$ via the inverse transform

$$h(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikx} H(k) dk. \quad (28)$$

Before this, it is instructive to consider two simpler but related problems in which there is no discontinuity in the boundary conditions.

2.1 The Reynolds Ridge

If the boundary condition (7) holds over the entire surface then we replace (27) with

$$H + l^2 k^2 H = \frac{U^2}{g} \frac{k^2}{\Gamma} (H - D). \quad (29)$$

We now define the Weber number m based on the length scale l

$$m = \frac{U^2}{gl} \quad (30)$$

²We have also assumed that $d \rightarrow 0$ as $|x| \rightarrow \infty$, which is not the case for a Blasius boundary layer. This difficulty can be resolved by introducing a slowly-decaying exponential factor in $d(x)$, then allowing the decay rate to tend to zero.

and the operator

$$\mathcal{D}_{\text{sea}}(k) = 1 + l^2 k^2 - mlk^2/\Gamma(k) \quad (31)$$

so that (29) may be written in the form

$$\mathcal{D}_{\text{sea}}(k)H(k) = -mlk^2 D(k)/\Gamma(k). \quad (32)$$

We note that $\mathcal{D}_{\text{sea}}(k) = 0$ is the dispersion relation for standing capillary-gravity waves on the free surface:

$$\mathcal{D}_{\text{sea}}(k) = 0 \quad (33)$$

$$\Leftrightarrow (Uk)^2 = g\Gamma(k)(1 + l^2 k^2). \quad (34)$$

The full dispersion relation is obtained by replacing $Uk \rightarrow -\omega + Uk$. The phase and group velocities for capillary-gravity waves are plotted in Figure 3. Provided that there

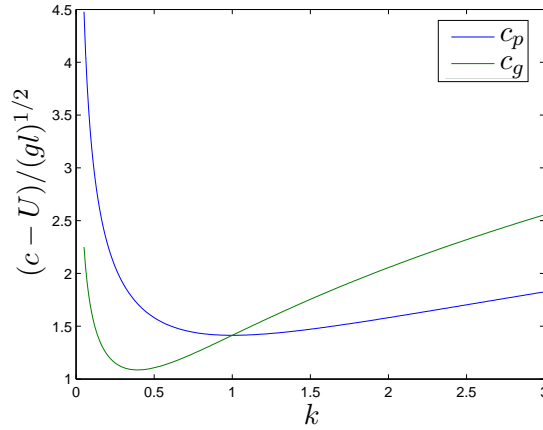


Figure 3: The phase speed c_p and group speed c_g of capillary-gravity waves in a flow of speed U .

are no roots of \mathcal{D}_{sea} for real k we may readily solve (32) for $H(k)$ and then find the free surface height from (28) by integrating along the real k axis, leading to the solution for the Reynolds ridge problem presented by Harper & Dixon. In general, $\mathcal{D}_{\text{sea}}(k)$ has four complex roots, which we label as k_{s1} , k_{s2} , k_{s3} and k_{s4} according to their location in the complex k -plane (see Figure 4). For $m > 2$ all four roots fall onto the real k axis, corresponding to the existence of standing waves; we must then deform the integration contour in (28) to avoid the singularities in the integrand. The particular choice of integration contour will determine which waves are present as $x \rightarrow \pm\infty$, and so we must apply a suitable radiation condition. We see from Figure 3 that the group speed exceeds the phase speed for short wavelength (capillary) waves, and the converse for long wavelength (gravity) waves. We should therefore deform our integration contour so that standing capillary waves appear upstream and gravity waves downstream in the solution [4]. For $x > 0$, the integrand in (28) is exponentially small for $\text{Im } k < 0$, so we can close our integration contour at infinity in the lower half k -plane; the solution will only contain waves arising from poles located in the \ominus domain. So we formally regard the two roots of \mathcal{D}_{sea} with smaller wavenumber k (corresponding to gravity waves) as residing in the lower half k -plane, and label them as k_{s3} and k_{s4} . This leads us to adopt the deformed integration contour shown in Figure 5.

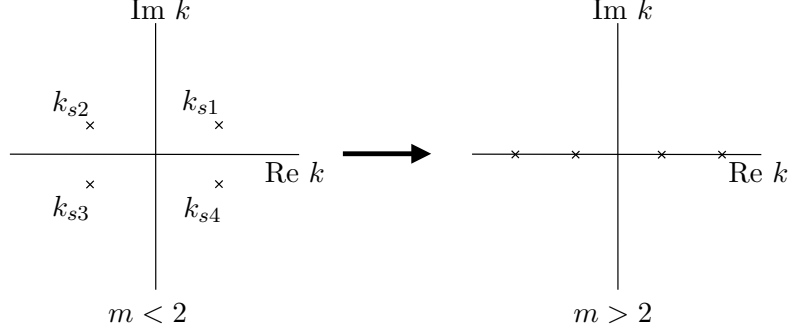


Figure 4: The roots of \mathcal{D}_{sea} fall onto the real line for $m > 2$.

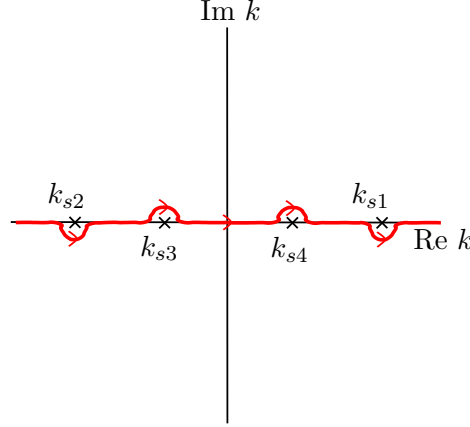


Figure 5: The integration contour used in the inverse transform (28) when $m > 2$.

2.2 Infinite Icesheet

Sufficiently far into the ice covered region we might expect effects from the ice edge to be negligible, and therefore apply boundary condition (5) for all x . In this case, (27) is replaced by

$$H + L^4 k^4 H = \frac{U^2 k^2}{g \Gamma} (H - D). \quad (35)$$

We proceed by analogy with §2.1; we define a Weber number

$$n = \frac{U^2}{gL} \quad (36)$$

and a dispersion operator

$$\mathcal{D}_{\text{ice}}(k) = 1 + L^4 k^4 - n L k^2 / \Gamma(k) \quad (37)$$

so that (35) may be written in the form

$$\mathcal{D}_{\text{ice}}(k) H(k) = -n L k^2 D(k) / \Gamma(k). \quad (38)$$

$\mathcal{D}_{\text{ice}}(k)$ also has four complex roots, which fall on the real axis when $n > (\frac{256}{27})^{1/4} \approx 1.75$; these roots correspond to standing flexural waves on the ice surface. As in the previous

section, we must deform our integration contour around these singularities when evaluating the inverse transform (28). Due to the similarities between $\mathcal{D}_{\text{ice}}(k)$ and $\mathcal{D}_{\text{sea}}(k)$, the correct contour is qualitatively identical to that sketched in Figure 5.

3 The Wiener–Hopf Technique

We begin by rewriting (27) in the form

$$\frac{\mathcal{D}_{\text{ice}}}{\mathcal{D}_{\text{sea}}}(k) = \frac{(1 + L^4 k^4)D - (L^4 k^4 - l^2 k^2)H_- - P(k)}{(1 + l^2 k^2)D + (L^4 k^4 - l^2 k^2)H_+ - P(k)} \quad (39)$$

To proceed with the Wiener–Hopf method we must split the LHS of (39) into a product of \oplus and \ominus functions:

$$\frac{\mathcal{D}_{\text{ice}}}{\mathcal{D}_{\text{sea}}}(k) = -\mathcal{K}_+(k)\mathcal{K}_-(k). \quad (40)$$

The functions \mathcal{K}_{\pm} must have neither poles nor zeros in their respective domains. We describe the splitting procedure in detail in §A; for now we simply observe that the split is unique up to multiplication of \mathcal{K}_{\pm} by constants, and that $\mathcal{K}_{\pm}(k) = O(k)$ as $|k| \rightarrow \infty$. We now deduce from (39) that

$$\mathcal{K}_+[(L^4 k^4 - l^2 k^2)H_+ + (1 + l^2 k^2)D - P] = \frac{1}{\mathcal{K}_-}[(L^4 k^4 - l^2 k^2)H_- - (1 + L^4 k^4)D + P(k)] \quad (41)$$

If $D(k)$ is an entire function then this expression is of the form

$$A_+(k) = A_-(k) \quad (42)$$

where A_+ and A_- are analytic in, respectively, the \oplus and \ominus regions of the complex k -plane. Therefore $A_{\pm}(k)$ can be extended to an entire function $A(k)$, which is necessarily a polynomial. If, however, $D(k)$ contains finite poles, then these must first be removed. For example, if d is given by

$$d(x, k_0) = \begin{cases} 0, & \text{for } x < 0 \\ 2 \sin(k_0 x), & \text{for } x > 0 \end{cases} \quad (43)$$

with $k_0 \in \mathbb{R}$ then

$$D(k, k_0) = D_+(k, k_0) = \frac{1}{k + k_0} - \frac{1}{k - k_0}. \quad (44)$$

Since this is a \oplus function, the simple poles at $k = \pm k_0$ must formally be regarded as residing in the lower half plane. We therefore define

$$\begin{aligned} A_+(k, k_0) &= \mathcal{K}_+[(L^4 k^4 - l^2 k^2)H_+ + \frac{1 + l^2 k^2}{k + k_0} - \frac{1 + l^2 k^2}{k - k_0} - P] \\ &\quad + \frac{1 + L^4 k_0^4}{\mathcal{K}_-(-k_0)(k + k_0)} - \frac{1 + L^4 k_0^4}{\mathcal{K}_-(k_0)(k - k_0)} \end{aligned} \quad (45)$$

$$\begin{aligned} A_-(k, k_0) &= \frac{1}{\mathcal{K}_-}[(L^4 k^4 - l^2 k^2)H_- - \frac{1 + L^4 k^4}{k + k_0} + \frac{1 + L^4 k^4}{k - k_0} + P] \\ &\quad + \frac{1 + L^4 k_0^4}{\mathcal{K}_-(-k_0)(k + k_0)} - \frac{1 + L^4 k_0^4}{\mathcal{K}_-(k_0)(k - k_0)} \end{aligned} \quad (46)$$

and conclude, as before, that $A_{\pm}(k, k_0) = A(k, k_0)$, a polynomial in k . The true boundary layer $d(x)$, given by (3), can be written as a superposition of modes of the form (43). In particular,

$$d(x) = \frac{1.7208}{2\sqrt{2\pi}} \left(\frac{\nu}{U} \right)^{1/2} \int_0^\infty \frac{d(x, k_0)}{k_0^{3/2}} dk_0 \quad (47)$$

The free surface $h(x)$ can likewise be written as a superposition of single mode solutions $h(x, k_0)$.

3.1 The single mode solution

We find from (45) and (46) that

$$H_+(k, k_0) = \frac{\frac{1}{\mathcal{K}_+} \left(A(k, k_0) + \frac{1+L^4 k_0^4}{\mathcal{K}_-(k_0)(k-k_0)} - \frac{1+L^4 k_0^4}{\mathcal{K}_-(-k_0)(k+k_0)} \right) - \frac{1+l^2 k^2}{k+k_0} + \frac{1+l^2 k^2}{k-k_0} + P}{L^4 k^4 - l^2 k^2} \quad (48)$$

$$H_-(k, k_0) = \frac{\mathcal{K}_- \left(A(k, k_0) + \frac{1+L^4 k_0^4}{\mathcal{K}_-(k_0)(k-k_0)} - \frac{1+L^4 k_0^4}{\mathcal{K}_-(-k_0)(k+k_0)} \right) + \frac{1+L^4 k^4}{k+k_0} - \frac{1+L^4 k^4}{k-k_0} - P}{L^4 k^4 - l^2 k^2} \quad (49)$$

$$\therefore H(k, k_0) = \frac{-2k_0}{k^2 - k_0^2} + \frac{\left(A(k, k_0) + \frac{1+L^4 k_0^4}{\mathcal{K}_-(k_0)(k-k_0)} - \frac{1+L^4 k_0^4}{\mathcal{K}_-(-k_0)(k+k_0)} \right) \left(\mathcal{K}_-(k) + \frac{1}{\mathcal{K}_+(k)} \right)}{L^4 k^4 - l^2 k^2} \quad (50)$$

We find from (50) that $H(k, k_0)$ is $O(A/k^3)$ as $k \rightarrow \infty$. In order for $h(x)$ to be continuous, $H(k)$ must be $O(1/k^2)$, and so $A(k, k_0)$ must therefore be $O(k)$; we write $A(k) = a(k_0) + b(k_0)k$. The constants a and b are determined by the condition that H_{\pm} should be analytic for $\text{Im } k \geq 0$, i.e. that the zeros in the denominators of (48) and (49) must represent removable singularities. Since $P(k)$ is $O(k^2)$ at $k = 0$, we therefore require that

$$A(k, k_0) + \frac{1+L^4 k_0^4}{\mathcal{K}_-(k_0)(k-k_0)} - \frac{1+L^4 k_0^4}{\mathcal{K}_-(-k_0)(k+k_0)} + \frac{1+L^4 k^4}{\mathcal{K}_-(k)(k+k_0)} - \frac{1+L^4 k^4}{\mathcal{K}_-(k)(k-k_0)} \quad (51)$$

is also $O(k^2)$. This implies that

$$k_0 a(k_0) = -\frac{2}{\mathcal{K}_-(0)} + \frac{1+L^4 k_0^4}{\mathcal{K}_-(k_0)} + \frac{1+L^4 k_0^4}{\mathcal{K}_-(-k_0)} \quad (52)$$

$$k_0^2 b(k_0) = -2k_0(1/\mathcal{K}_-)'|_0 + \frac{1+L^4 k_0^4}{\mathcal{K}_-(k_0)} - \frac{1+L^4 k_0^4}{\mathcal{K}_-(-k_0)} \quad (53)$$

For notational convenience, we now define

$$\tilde{A}(k, k_0) \equiv A(k, k_0) + \frac{1+L^4 k_0^4}{\mathcal{K}_-(k_0)(k-k_0)} - \frac{1+L^4 k_0^4}{\mathcal{K}_-(-k_0)(k+k_0)} \quad (54)$$

$$= -\frac{2}{k_0} \left((1/\mathcal{K}_-)|_0 + (1/\mathcal{K}_-)'|_0 k \right) + \frac{k^2(1+L^4 k_0^4)}{k_0^2 \mathcal{K}_-(k_0)(k-k_0)} - \frac{k^2(1+L^4 k_0^4)}{k_0^2 \mathcal{K}_-(-k_0)(k+k_0)} \quad (55)$$

We can now find the free surface height through the inverse transformation (28). In order to evaluate this integral we make use of the fact that

$$\frac{\mathcal{K}_- + 1/\mathcal{K}_+}{L^4 k^4 - l^2 k^2} = \frac{\mathcal{K}_-}{\mathcal{D}_{\text{ice}}} \quad (56)$$

$$= \frac{-1}{\mathcal{K}_+ \mathcal{D}_{\text{sea}}} \quad (57)$$

For $x > 0$, the integrand in (28) decays at infinity for $\text{Im } k < 0$. We therefore apply the identity (56) and rewrite equation (50) as

$$H(k, k_0) = \frac{-2k_0}{k^2 - k_0^2} + \tilde{A}(k, k_0) \frac{\mathcal{K}_-(k)}{\mathcal{D}_{\text{ice}}(k)} \quad (58)$$

We calculate $h(x, k_0)$ by collapsing the integration contour in (28) around the branch cut in the lower half plane. In doing so, we pick up pole contributions from the poles at k_{i3} , k_{i4} and $\pm k_0$, all of which are defined as residing in the lower half plane. The solution is

$$\begin{aligned} h(x, k_0) = & -\frac{2nL|k_0| \sin(k_0 x)}{1 + L^4 k_0^4 - nL|k_0|} + 2 \text{Im} \left\{ \frac{\tilde{A}(k_{i4}, k_0) \mathcal{K}_-(k_{i4})}{4L^4 k_{i4}^3 - nL} e^{-ik_{i4} x} \right\} \\ & - \frac{1}{\pi} \int_0^\infty d\lambda \frac{nL\lambda \tilde{A}(-i\lambda, k_0) \mathcal{K}_-(-i\lambda)}{(1 + L^4 \lambda^4)^2 + n^2 L^2 \lambda^2} e^{-\lambda x} \end{aligned} \quad (59)$$

The corresponding solution for $x < 0$ is

$$\begin{aligned} h(x, k_0) = & 2 \text{Im} \left\{ \frac{\tilde{A}(k_{s1}, k_0) / \mathcal{K}_+(k_{s1})}{2l^2 k_{s1} - nL} e^{-ik_{s1} x} \right\} \\ & + \frac{1}{\pi} \int_0^\infty d\lambda \frac{nL\lambda \tilde{A}(i\lambda, k_0) / \mathcal{K}_+(i\lambda)}{(1 - l^2 \lambda^2)^2 + n^2 L^2 \lambda^2} e^{\lambda x} \end{aligned} \quad (60)$$

3.2 The full solution

We deduce from (47) that the free surface height is

$$h(x) = \frac{1.7208}{2\sqrt{2\pi}} \left(\frac{\nu}{U} \right)^{1/2} \int_0^\infty \frac{h(x, k_0)}{k_0^{3/2}} dk_0 \quad (61)$$

Since \tilde{A} is $O(k_0)$ as $k_0 \rightarrow 0$ (which can be deduced from (55)), the obvious singularity in (61) is integrable, and the solution is well defined.

4 Results

The effect of the ice edge can be seen by comparing (58) with the solution for an infinite icesheet (38). Indeed we can recover (38) by replacing

$$\tilde{A} \mathcal{K}_- \longrightarrow -(1 + L^4 k^4) D \quad (62)$$

in (58). In Figure 6 we illustrate the difference in a particular case. We find, as we might expect on physical grounds, that the presence of the ice edge significantly alters the small scale structure of the solution, but has negligible influence in the limit of zero wavenumber. Both sides of (62) are $-2/k_0 + O(k^2)$ as $k \rightarrow 0$.

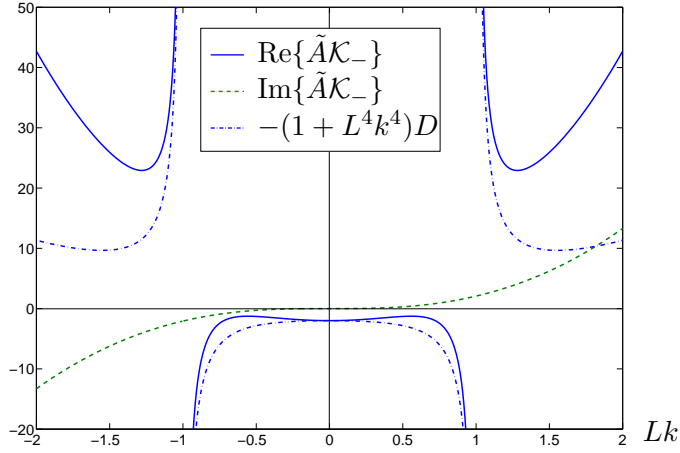


Figure 6: The effect of the ice edge on the single mode solution in the case $k_0 = 1/L$, $n = 1.5$, $l/L = 0.01$. The real and imaginary parts of $\tilde{A}(k, k_0)\mathcal{K}_-(k)$ are shown along with the function $-(1 + L^4 k^4)D(k, k_0)$.

A Wiener–Hopf Splitting

We need to find $\mathcal{K}_\pm(k)$ such that

$$\frac{\mathcal{D}_{\text{ice}}}{\mathcal{D}_{\text{sea}}}(k) = -\mathcal{K}_+(k)\mathcal{K}_-(k). \quad (63)$$

To this end, we define

$$\mathcal{R}_{\text{ice}} = \frac{\mathcal{D}_{\text{ice}}(k)}{L^4(k - k_{i1})(k - k_{i2})(k - k_{i3})(k - k_{i4})} \quad (64)$$

$$\mathcal{R}_{\text{sea}} = \frac{\mathcal{D}_{\text{sea}}(k)\Gamma(k)^2}{l^2(k - k_{s1})(k - k_{s2})(k - k_{s3})(k - k_{s4})} \quad (65)$$

Here, k_{ij} and k_{sj} represent the roots of \mathcal{D}_{ice} and \mathcal{D}_{sea} lying in the j -th quadrant of the complex k plane. We note that $k_{s2} = -k_{s1}^*$ and $k_{s4} = -k_{s3}^*$ (similarly for the roots of \mathcal{D}_{ice}). The functions \mathcal{R}_{ice} and \mathcal{R}_{sea} have only branch cuts in their analytic structure, and tend to unity for large k . Thus we may define

$$\log \mathcal{R}_{\text{ice}+}(k) = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \log[\mathcal{R}_{\text{ice}}(k')] \frac{dk'}{k' - k} \quad (66)$$

$$\log \mathcal{R}_{\text{sea}+}(k) = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \log[\mathcal{R}_{\text{sea}}(k')] \frac{dk'}{k' - k} \quad (67)$$

$$(68)$$

where the integration contour lies below the point $k' = k$. The integrands are continuous provided that we take the branch of the logarithm with a cut along the negative real axis.

The integration contour may then be collapsed around the branch cut in the lower half k' plane, which leads to

$$\log \mathcal{R}_{\text{ice}\pm}(k) = \frac{1}{\pi} \int_0^\infty \tan^{-1} \left[\frac{n\lambda}{1+\lambda^4} \right] \frac{d\lambda}{\lambda \mp iLk} \quad (69)$$

$$\log \mathcal{R}_{\text{sea}\pm}(k) = \frac{1}{\pi} \int_0^\infty \tan^{-1} \left[\frac{n\lambda}{1-l^2\lambda^2/L^2} \right] \frac{d\lambda}{\lambda \mp iLk} \quad (70)$$

Some care is required in choosing the appropriate branches of \tan^{-1} in these expressions. For the first, we require $\tan^{-1} \in [0, \pi)$, but for the second, $\tan^{-1} \in (-\pi, 0]$. This means that the integrand in (70) is $\sim \frac{-1}{\lambda \mp iLk}$ as $\lambda \rightarrow 0$, implying a logarithmic singularity as $k \rightarrow 0$. We may now write

$$\mathcal{K}_+(k) = \frac{L^2(k - k_{i3})(k - k_{i4})}{l(k - k_{s3})(k - k_{s4})} k \frac{\mathcal{R}_{\text{ice}+}(k)}{\mathcal{R}_{\text{sea}+}(k)} \quad (71)$$

$$\mathcal{K}_-(k) = -\frac{L^2(k - k_{i1})(k - k_{i2})}{l(k - k_{s1})(k - k_{s2})} k \frac{\mathcal{R}_{\text{ice}-}(k)}{\mathcal{R}_{\text{sea}-}(k)} \quad (72)$$

We note that $\mathcal{K}_\pm^*(k) = -\mathcal{K}_\pm(-k^*)$ and that $\mathcal{K}_\pm(k) = O(k)$ as $|k| \rightarrow \infty$.

References

- [1] N. J. BALMFORTH AND R. V. CRASTER, *Ocean waves and ice sheets*, Journal of Fluid Mechanics, 395 (1999), pp. 89–124.
- [2] S. GOLDSTEIN, *Modern developments in fluid dynamics*, Oxford University Press, (1938).
- [3] J. F. HARPER AND J. N. DIXON, *The leading edge of a surface film on contaminated flowing water*, Fifth Australasian Conference on Hydraulics and Fluid Mechanics, (1974).
- [4] H. LAMB, *Hydrodynamics*, Dover Publications, (1932).
- [5] R. D. MINDLIN, *Influence of rotatory inertia and shear on flexural motions of isotropic, elastic plates*, Journal of Applied Mechanics, 18 (1951), pp. 31–38.

Equilibria of diffusive moist static energy balance model

Yutian Wu

January 12, 2009

Introduction

Temperature is a key variable in the measurement of climate. The first order approximation of temperature can be derived from the global energy balance model. The basic idea of the global energy balance of the Earth is that the incoming solar radiation should always keep in balance with the energy back to space by Earth's radiative emission. Because of the transparency of the Earth's atmosphere to the incoming solar radiation and the opaque characteristics to the terrestrial radiation, the absorption of solar radiation mostly takes place at the surface while most of the emission happens at top of the atmosphere. Budyko (1969)[4] and Sellers (1969)[11] brought up a type of energy balance model by using the zonal and annual averaging but allowing the latitudinal dependences of surface temperature, albedo and meridional heat transport. One of the major features of this type of model is that it presented the possibility of an abrupt transition to a completely ice-covered Earth if the solar constant is lowered by only a few percent.

My project aims to figure out the effect of water vapor in this type of energy balance model. The effect of water vapor is significant on climate but complicated to analyze. First of all, water vapor is one of the most important greenhouse gases of the atmosphere and has a positive feedback on climate change: if anything happens to initiate a small temperature increase of the atmosphere, due to the saturation vapor pressure given in Clausius-Clapeyron equation, more water vapor can be held in the atmosphere. The additional water vapor will lead to additional greenhouse effect which warms the Earth further beyond the initial warming. Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4)[7] suggested that water vapor changes represent the largest positive feedback affecting equilibrium climate sensitivity, and it is this key positive feedback that amplifies the response to the anthropogenic forcing. In addition, water vapor is playing a major role in determining the atmospheric convection and cloud cover, which in turn exerts complicated feedback on climate. However, cloud feedbacks still remain the largest source of uncertainty in the projection of climate change. Finally water vapor is an important component in atmospheric circulation and hydrological cycles. With increased moisture in the atmosphere in the future climate baroclinic eddies may increase in strength, on the other hand, the temperature gradient from the equator to the poles may decrease due to the increased poleward moisture energy transport. Because of the many aspects of the impacts of water vapor on the climate and various complicated water vapor feedbacks, the net effect of water

vapor is complicated to analyze in real climate. In this study, we develop simple numerical models with parameterizations of basic but essential physics. Rather than full-blown General Circulation Models, simplified models can be understood completely and may be able to address big questions in real climate. Section 1 introduces an one-layer diffusive moist static energy balance model. Rather than sensible heat merely, moist static energy is applied which includes not only sensible heat but also latent heat and potential energy. Although the one-layer model has represented several interesting features, it cannot produce modern climate even though all the model parameters are of modern values; in addition, not much water vapor effect can be realized except the substantial atmospheric moisture transport to the poles. The reason for it might be the one-layer model itself is too simple to represent the real situation of the climate, for example clouds, which is an important component of climate associated sensitively with water vapor in the atmosphere. Clouds have substantial interactions with both solar and terrestrial radiation in both radiation reflection and absorption, therefore a net energy loss or gain to the atmosphere-Earth system can be produced by clouds depending on their characteristics. Although the net nonradiative and radiative effect of clouds is hard to predict, clouds do have the potential for climate change. Lots of study has been focused on the physics of clouds and narrowing down the uncertainty of cloud-climate feedback in model simulations. In Section 2 a simple two-layer atmosphere-surface model is developed to implement the most basic but essential physics of one of the most important features of water vapor, i.e. cloud feedback. In the discussion we apply the basic idea of convective cloud feedback which was first proposed by Abbot and Tziperman (2007)[2] to explain the equable climate during the late Cretaceous and early Paleogene (about 100 million to 35 million years ago). Geological data suggests that the Earth was much warmer than it is today, especially at high latitudes and during winter time. For example, there were palm trees in Wyoming and crocodiles on islands in the Arctic ocean. The equator to pole temperature difference and the amplitude of the seasonal cycle at high latitudes were relatively small, hence the name "equable climate". The basic idea of convective cloud feedback is that an initial warming, for example, due to the increase of CO_2 concentration, leads to the moist destabilization of the high-latitude atmosphere and initiates atmospheric convection. The resulting convective clouds and atmospheric moisture both help trap the outgoing longwave radiation, leading to further warming. Furthermore since convective cloud feedback allows for multiple equilibria[2], more equilibria are expected in the two-layer model when compared with the equilibria in the one-layer one. The physics of those equilibria is also interesting to know, which is demonstrated in Section 3. In Section 4, a seasonal model is set up when both the time-varying solar insolation and cloud albedo effect are taken into account, for example, more solar incoming radiation is reflected when convection happens due to the cloud albedo effect. We also explore the numerical time-dependent solutions of the seasonal model.

One-layer Model

A one-layer model is set up based on the Budyko-Sellers type of energy balance model, in which the basic idea of global energy balance is applied and the energy transport is diffusive.

The model is defined by [8][9][5]

$$C \frac{\partial T}{\partial t} = \frac{\tilde{D}}{a^2 \cos(\phi)} \frac{\partial}{\partial \phi} [\cos(\phi) \frac{\partial M}{\partial \phi}] + S - OLR \quad (1)$$

where C is the thermal inertia, T is the temperature and \tilde{D} is the diffusion coefficient. a is the Earth radius, ϕ is the latitude. M is the moist static energy which is defined as $M = C_p T + Lq + gz$, where C_p is the specific heat capacity at constant pressure, L is the latent heat release, q is the specific humidity and z is the height of the parcel. S is the incoming solar radiation which is given by

$$S = L_\odot (1 - \alpha) F(\phi) \quad (2)$$

where $L_\odot = 1370 \text{ W/m}^2$ is the solar constant. α is the albedo. Albedo is not a static quantity, however, is determined sensitively by the processes in the atmosphere and at the surface. As climate changes, the surface characteristics change, for example, the proportion of the Earth's surface covered by ocean water, sea ice, land ice, vegetation and so on, and the resulting albedo changes feed back on the state of the climate. In this model I choose a simple but reasonable representation of albedo which is given by

$$\alpha(T) = \begin{cases} \alpha_i & \text{for } T \leq T_i \\ \alpha_o + (\alpha_i - \alpha_o) \frac{(T - T_o)^2}{(T_i - T_o)^2} & \text{for } T_i < T < T_o \\ \alpha_o & \text{for } T \geq T_o \end{cases} \quad (3)$$

where α_i and α_o are albedos of sea ice and ocean water, respectively, and T_i and T_o are two threshold temperatures. When surface temperature is very large within some area, i.e. $T_s > T_o$, it is above freezing and there is no ice, thus in this case, the albedo is given by the ocean water albedo; when surface temperature is very low, i.e. $T_s < T_i$, it is completely covered by sea ice thus the albedo is given by the sea ice albedo within some area. In between it is reasonable to interpolate the albedo by a quadratic form given above. F is the annual mean flux factor of solar incoming radiation, which gives the latitudinal distribution of solar flux. It peaks at the equator and decreases gradually to the poles. OLR is the outgoing longwave radiation which is a function of the emission temperature T_e at top of the atmosphere, i.e. $OLR(T_e) = \sigma T_e^4$, where σ is the Stefan-Boltzmann constant. Budyko (1969)[4] applied a linearized version of OLR , based on the fact that T_e is linearly related to the surface temperature T_s , which is given by

$$OLR(T_e) \approx A + BT_s \quad (4)$$

where A and B are constants that can be deduced from the fitting of observations. A is roughly linear in $1/\log(CO_2)$ concentration. By analyzing the data of monthly mean upward longwave radiation flux and surface air temperature from year 1961 to 2000 in NCEP/NCAR Reanalysis Data, $A = 211.2 \text{ W/m}^2$ and $B = 2.09 \text{ W/m}^2(^{\circ}\text{C})^{-1}$ produces the best fit between the annual mean fluxes calculated using the linearized approximation and the reanalysis data.

By defining variable $y = \sin(\phi)$, the equation of steady-state solution can be rewritten as

$$\frac{\tilde{D}}{a^2} \frac{d}{dy} [(1 - y^2) \frac{dM}{dy}] = OLR(T_s(y)) - S(y). \quad (5)$$

Steady-state solutions are important because a Lyapunov functional has been found in the previous PDE which implies all time-dependent solutions will eventually converge to stable steady-state solutions just like the case in Budyko-Sellers type of energy balance model. Because it is symmetric between the Northern and Southern Hemisphere in the model, only northern hemisphere is discussed here, i.e. $0 \leq y \leq 1$. For boundary conditions, take $\frac{dT}{dy} = 0$ at the equator, which enforces symmetry between the two hemispheres, and insist that T be regular at the poles.

Since M is conserved under conditions of either dry adiabatic ascent (exactly) or saturated adiabatic ascent (approximately), M can be expressed in the term of T_s based on the surface moist static energy and Clausius-Clapeyron equation, i.e.

$$M = C_p T + Lq + gz \quad (6)$$

$$= C_p T_s + L_v q_s \quad (7)$$

$$= C_p T_s + h_s L_v q_s^* \quad (8)$$

$$= C_p T_s + h_s L_v \frac{R_d}{R_v} \frac{1}{p_s} e_s^* \quad (9)$$

$$= C_p T_s + h_s L_v \frac{R_d}{R_v} \frac{1}{p_s} e_s(T_o) \exp\left[\frac{L_v}{R_v} \left(\frac{1}{T_o} - \frac{1}{T_s}\right)\right] \quad (10)$$

where the subscripts s represent surface values, the superscripts $*$ represent saturation values. L_v is the latent heat release of evaporation, h_s is the surface relative humidity which is assumed to be a constant everywhere, q_s^* is the saturation specific humidity at the surface, $T_o = 273.15K$, R_d and R_v are the gas constants for dry air and water vapor, respectively, p_s is the surface pressure, and $e_s(T_o) = 6.11mb$ is the saturation vapor pressure at T_o .

Now the climate is characterized by the annual average and global mean surface temperature T_s , and the steady-state equation becomes an nonlinear ODE of T_s which can be solved numerically. For fixed A and B (fixed CO_2 concentration and thus fixed OLR in this model), for example, $A = 211.2W/m^2$ and $B = 2.09W/m^2(^{\circ}C)^{-1}$ which are the parameters derived from fitting the current climate data, but for different initial guesses, steady-state solutions are not always the same, which implies the existence of bifurcation. Hence instead I solve A (for fixed B) as an unknown parameter by prescribing the equatorial surface temperature $T_s(y = 0)$. Figure 1 shows the steady-state solutions of this model as a function of A . In particular, not only equatorial surface temperature $T_s(y = 0)$ but also polar surface temperature $T_s(y = 1)$ and mean surface temperature are drawn. The bifurcation diagram is composed of a stable ice-free climate a , an unstable climate b and a stable snowball climate c . In particular, for $A = 211.2W/m^2$, three possible climate states are shown in Figure 2. They are a stable snowball climate where temperature is below freezing everywhere on the Earth, an unstable climate which is $20K$ lower than the current climate everywhere from the equator to the poles, and a stable ice-free climate where temperature is relatively high everywhere and the temperature difference between the equator and poles is relatively small, respectively. As for the unstable climate, although it is $20K$ lower than the current climate, the temperature gradient from the equator to the poles which is about $40K$ is quite reasonable. In addition, a comparison of the steady-state solutions has been done between the dry (Figure 3) and moist diffusive energy balance model. For the dry

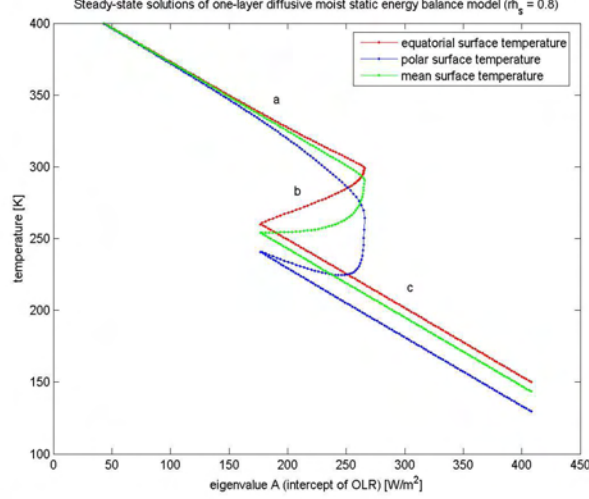


Figure 1: Steady-state solutions of one-layer diffusive moist static energy balance model as a function of A , which is roughly linear in $1/\log(CO_2)$ concentration. Relative humidity is assumed to be 0.8 everywhere. Model parameters are: $\alpha_i = 0.6, \alpha_o = 0.1, T_i = 260K, T_o = 290K, B = 2.09W/m^2(^{\circ}C)^{-1}$.

diffusive energy balance model, there is no latent heat in the energy term but only sensible heat. The major difference between the steady-state solutions of the two models is that for the ice-free climate the temperature gradient from the equator to the poles is relatively small in the moist case while it is roughly $50K$ in the dry one. My explanation for such a difference is that in hot climate water vapor helps transport large amount of energy poleward effectively and thus wipes out the temperature gradient. This is a distinguished role that water vapor is playing in this one-layer model.

It is also noticed that by building up (or reducing) the atmospheric CO_2 concentration, the disappearance of the snowball climate (or ice-free climate) will be induced. For example, based on this model, when $A < 177W/m^2$, which implies a very high CO_2 concentration, the stable snowball climate will disappear; when $A > 266W/m^2$, which implies a very low CO_2 concentration, the stable ice-free climate will disappear. Although these critical values might be too rough to make any certain conclusions, it presents a simple trend of climate states as the atmospheric CO_2 concentration changes.

Therefore basically three steady states have been found in this one-layer diffusive moist static energy balance model, of which one represents a stable ice-free climate where temperature is relatively high everywhere and the temperature gradient is relatively small from the equator to the poles. The other one shows an unstable climate while the third one is a snowball climate. In addition, this model represents the possibility of abrupt climate transitions, for example, the transition from the snowball climate to the unstable climate state, or and the transition from the unstable climate state to the ice-free climate state when temperature gets higher due to increased atmospheric CO_2 concentration, for example. Finally the model demonstrates one of the most important roles that water vapor is playing on the real climate, i.e. the meridional moisture energy transport that helps wipe

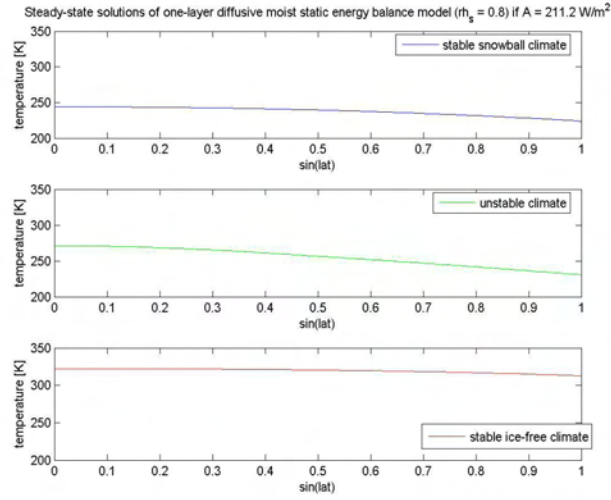


Figure 2: Steady-state solutions of one-layer diffusive moist static energy balance model if $A = 211.2 \text{ W/m}^2$. The top, middle and bottom panels present the global mean surface temperature in a stable snowball climate, an unstable climate which is 20K lower than the current one, and a stable ice-free climate, respectively. Model parameters are the same as before.

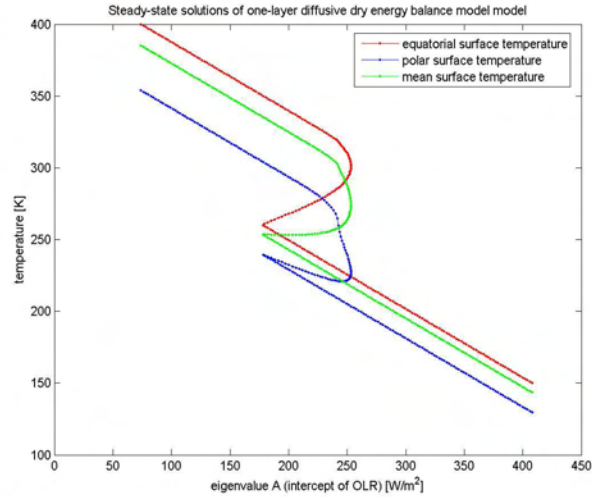


Figure 3: Steady-state solutions of one-layer diffusive dry energy balance model as a function of A . Same model parameters.

out the temperature gradient from the equator to the poles. However, the model has several limitations: one is that it cannot produce the current climate, either the temperature of the possible states is too high or too low. The second existing problem is that the effect of water vapor can not be much realized in this simple model.

Since this model is of one-layer, many important effects of water vapor can not be parameterized, such as the greenhouse effect, atmospheric convection and cloud cover. Hence we need to modify the model by adding some representation of cloud effects and hope for better outputs, such as the simulation of modern climate and so on.

Two-Layer Model

a. Developing the Model

An important effect of water vapor on the climate is the formation of clouds which have substantial interactions with both solar and terrestrial radiation. However, the impact of an individual cloud on the local energy balance depends sensitively on its physical characteristics, for example, the total mass of water, the size and shape of the droplets or particles and so on. Generally speaking, low-altitude stratus clouds have a large and negative cloud radiative forcing while high-altitude clouds are associated with a neutral or positive cloud radiative forcing. For my further discussion, I apply a simple representation of the greenhouse effect of clouds at first. As for the cloud albedo effect, it is closely related to the solar flux at certain latitude and time. The model should have seasonal cycle included then since, for example, at polar darkness, there is only greenhouse effect of clouds and it is discussed in Section 4.

Abbot and Tziperman (2007)[2] proposed an idea of convective cloud feedback to help explain the equable climate during the late Cretaceous and early Paleogene (about 100 million to 35 million years ago). At high latitudes under lower solar insolation, it is potentially possible to have a positive cloud radiative forcing with clouds associated with convection. If the extratropical surface temperature had increased large enough to initiate high convection, clouds would change from low-altitude to high-altitude ones and thus cloud radiative forcing would switch to positive from negative which could lead to further warming and even more atmospheric convection. Simulations done by Abbot and Tziperman[3][2] have shown that not only the high temperature situation at high latitudes and during winter time but also the relative low equator to pole temperature difference can be reproduced with the idea of high-latitude convective cloud feedback.

The model is composed of two layers[1]. The top one represents the free troposphere (200 – 900mb) and the bottom one represents the surface layer which is composed of the mixed-layer ocean (top 50m) and the boundary layer of the atmosphere (900 – 1000mb) dominated by turbulence. The energy balance of this model can be written as

$$C_s \frac{\partial T_s}{\partial t} = \frac{\partial}{\partial y} [(1 - y^2) D_1 \frac{\partial T_s}{\partial y}] + L_{\odot} (1 - \alpha) F(y) - F_c + \epsilon \sigma T_a^4 - \sigma T_s^4 \quad (11)$$

$$C_a \frac{\partial T_a}{\partial t} = \frac{\partial}{\partial y} [(1 - y^2) D_2 \frac{\partial T_a}{\partial y}] + F_c + \epsilon \sigma (T_s^4 - 2T_a^4) \quad (12)$$

where C_s and C_a are the thermal inertia for the bottom and top layer, respectively; T_s and T_a are the surface and atmospheric temperature, respectively; D_1 and D_2 are the diffusive

coefficients for the bottom and top layer, respectively (the two diffusive coefficients should be different in principle but are assumed to be the same at first); the first terms on the right hand side of the equations are the diffusive heat transport at the bottom and top layers, respectively; F_c is the convective flux from the boundary layer to the free troposphere; ϵ is the emissivity of the free troposphere; and σ is the Stefan-Boltzmann constant. Boundary conditions of this two-layer model are similar to that of the one-layer model, i.e.

$$\frac{dT_s}{dy}(y=0) = \frac{dT_a}{dy}(y=0) = 0 \quad (13)$$

$$T_s(y=1) \text{ and } T_a(y=1) \text{ are regular.} \quad (14)$$

The value of convective flux F_c and emissivity ϵ depends on whether there is convection or not, which is in turn described by moist instability. We determine the moist instability by comparing the surface moist static energy

$$M_s = C_p T_s + L_v q_s \quad (15)$$

with the atmospheric saturation moist static energy

$$M_a^* = C_p T_a + L_v q_a^* + g z_a \quad (16)$$

where q_a^* is the atmospheric saturation specific humidity, g is the acceleration of gravity, z_a and p_a are the height and pressure of the atmospheric layer, respectively. The convective flux F_c and emissivity ϵ are given by

$$F_c = \begin{cases} \gamma(M_s - M_a^*) & \text{if } M_s \geq M_a^* \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

$$\epsilon = \begin{cases} \epsilon_o + \Delta\epsilon & \text{if } M_s \geq M_a^* \\ \epsilon_o & \text{otherwise} \end{cases} \quad (18)$$

If $M_s < M_a^*$, ascending parcel is stable to moist convection and thus no convection happens. In this nonconvecting case, F_c is set to be 0 and atmospheric emissivity ϵ is given by the background emissivity ϵ_o which is mainly determined by the atmospheric CO_2 concentration and is roughly linear in $\log(CO_2)$ (Sasamori, 1968 [10]). As long as the moist static energy of the ascending parcel reaches the atmospheric saturation moist static energy, i.e. $M_s = M_a^*$, the parcel becomes unstable to convection and releases latent heat to the surrounding environment, and thus F_c is set to be linear of $M_s - M_a^*$ with a linear coefficient γ given by $\frac{C_s}{C_p} \frac{1}{\tau}$, where τ is the time scale of convection, for example, about 1 – 10 hours, and atmospheric emissivity is increased by a certain amount $\Delta\epsilon$ due to convective clouds and increased atmospheric moisture.

b. Equilibria in the Model

The two joint nonlinear ODEs (steady state of equations 11 and 12) are solved numerically to find out the solutions of steady-state surface and atmospheric temperature. For the same background emissivity ϵ_o , steady-state solutions are sensitive to the initial guesses and thus there again exists bifurcation in the solutions. Same as the numerical method

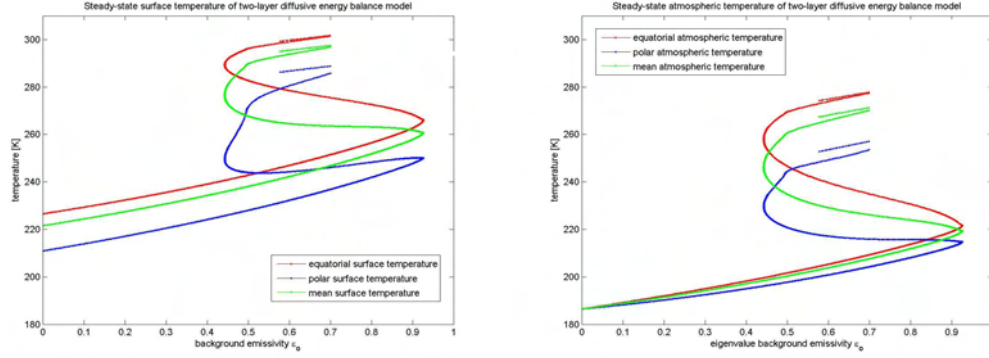


Figure 4: Steady-state surface temperature (left) and atmospheric temperature (right) of two-layer diffusive moist static energy balance model as a function of background emissivity ϵ_o , which is roughly linear in $\log(CO_2)$ concentration. Part of the convecting branch where the total atmospheric emissivity ϵ is larger than 1 is not physical and not shown here. Model parameters are: $\Delta\epsilon = 0.3$, $\gamma = 5.8101 kg/m^2/s$, $RH_s = 0.8$, $z_a = 5km$, $\alpha_i = 0.6$, $\alpha_o = 0.1$, $T_i = 260K$, $T_o = 290K$.

in one-layer model, we solve the background emissivity ϵ_o as an unknown eigenvalue by prescribing the equatorial surface temperature. Figure 4 shows the bifurcation diagrams of the steady-state surface and atmospheric temperature produced by the two-layer model. In particular, $\Delta\epsilon$ is set to be 0.3, and thus part of the bifurcation diagram where the total atmospheric emissivity ϵ is larger than 1 (or the background emissivity ϵ_o is larger than 0.7) has no physical meaning and is not drawn here since $0 \leq \epsilon \leq 1$. If it starts with a background emissivity larger than 0.7 and has convection happening, then the steady-state solution will stay with the atmospheric emissivity equals 1.

One of the interesting features of the bifurcation diagram of this model is that there are two more equilibrium solutions around $\epsilon_o = 0.6$ and when surface temperature is relatively high. Physical interpretation of the solutions in the following demonstrates that one is a stable equilibrium corresponding to a fully convecting state where surface temperature is high and convection happens everywhere from the equator to the poles, and the other one is an unstable equilibrium which is not shown in the figures. The problem with the unstable equilibrium branch (not shown) is that it is hard to find the solution numerically with the method other steady-state solutions are figured out. By doing projection onto Legendre polynomials and characterizing the steady-state solutions by first few modes, the unstable equilibrium branch can be found numerically but without high accuracy. Since the numerical computation is not reliable so far, it is not shown here.

In order to further understand the physics of the bifurcation diagram, a diagnosis of temperature and convection situation has been done in the following for various steady states, for example, states $< 1 > - < 11 >$ marked in Figure 5. Figure 6, 7, 8, 9 show the steady-state atmospheric and surface temperature, i.e. T_a and T_s , and $\Delta M = M_s - M_a^*$ which is a measure of whether there is convection or not as a function of y for these 11 states, respectively. If the surface temperature is very low, there is no possibility of convection where the moist instability is negative everywhere from the equator to the poles.

It is the same stable snowball climate as represented in one-layer model (states $< 1 >$ $< 2 >$ $< 3 >$ shown in Figure 6). If the equatorial surface temperature increases and reaches some critical value such that $M_s = M_a^*$, moist convection will firstly take place at the equator. The transition of convection at the equator corresponds to the first edge in the bifurcation diagram. As the equatorial surface temperature gets higher and higher, the upper atmospheric layer becomes moist unstable and convection can be initiated gradually from latitude to latitude. As shown in Figure 7 and 8, moist convection extends from the equator to the poles as the equatorial surface temperature increases, but since up to, for example, state $< 8 >$, convection doesn't come out at the poles yet, we name it partially convecting state. However, stability analysis tells that the branch where states $< 4 >$ and $< 5 >$ (in Figure 7) lie is unstable while the one which has states $< 6 >$, $< 7 >$, $< 8 >$ (in Figure 8) is stable, and this transition corresponds to the second edge of the bifurcation diagram. This is also consistent with the stability results in one-layer model. Eventually moist convection occurs at the poles when the surface equatorial temperature is large enough and this is what Figure 9 shows in which moist instability is positive at the poles. Due to the fact that convection takes place everywhere from the equator to the poles in states $< 9 >$, $< 10 >$, $< 11 >$, we name them fully convecting states. Although the unstable missing branch is hard to compute numerically, it is related to the transition of polar convection when the moist instability of state $< 8 >$ is compared with that of state $< 9 >$. The stability of each branch can also be checked by identifying the eigenvalues of the Jacobian matrix estimated at equilibrium states in perturbation equation. Therefore in conclusion the bifurcation diagram of the two-layer model is composed of several steady-state branches, for example, a stable nonconvecting branch, an unstable partially convecting branch, a stable partially convecting branch, an unstable branch which is related to the onset of polar convection, and a stable fully convecting branch.

Moreover the stable partially convecting state where the background emissivity $\epsilon_o = 0.46$ roughly represents the modern climate with $293K$ at the equator and $257K$ at the poles. It demonstrates that the implement of convective cloud feedback helps better simulate the real climate and the impact of clouds is significant on climate.

c. Climate Sensitivity in the Model

There are several free parameters in the two-layer model, such as surface relative humidity RH_s , height of the atmospheric layer z_a and linear coefficient of convective flux γ . Our next step is to study the climate sensitivity of this model by varying these free parameters. Climate sensitivity is a significant component and is defined as the relationship between the forcing and the magnitude of the climate change response. The study of climate sensitivity can help better understand and interpret the model and also the physics of it.

As the surface relative humidity increases, the parcel lifted upward from the surface can get saturated more easily and thus convection can be initiated more quickly. With a high surface relative humidity convection can occur early before the surface temperature there reaches a high value. The numerical comparison of the bifurcation diagrams of different surface relative humidity is shown in Figure 10. Figure 11 represents the steady-state solutions of different heights of atmospheric layer. For the climate sensitivity of the height of the atmospheric layer, as the height decreases, the onset of convection can be more easily

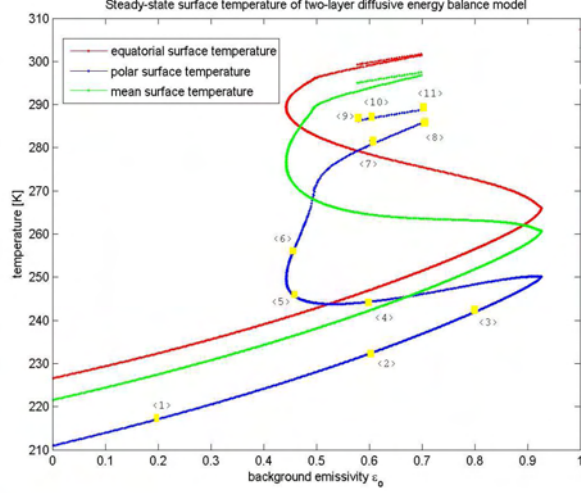


Figure 5: Steady-state surface temperature of two-layer diffusive moist static energy balance model as a function of background emissivity ϵ_o which is roughly linear in $\log(CO_2)$ concentration. Various steady-state solutions, for example, states $\langle 1 \rangle - \langle 11 \rangle$ are marked.

initiated at each latitude and the transition to both partially and fully convecting branches can be reached with a lower surface temperature compared with the case of a higher atmospheric layer. Figure 12 represents the steady-state solutions of different linear coefficients of convective flux. The linear coefficient of convective flux γ determines the magnitude of convective flux transferred from the boundary layer to the free troposphere when convection happens. Since the time scale of convection can vary within 1 – 10 hours, small changes of γ in this model are physically reasonable. As γ gets smaller, the top atmospheric layer gets less convective flux from the bottom one, which reduces the saturation moist static energy in the atmospheric layer. In this case, convection again can be more easily initiated which leads to the situation of fully convection at a lower surface temperature compared with the case with larger γ . The difference between this case and the former two is that varying the value of γ only changes the critical surface temperature of the onset of moist convection but keeps the structure (relationship between background emissivity and surface temperature) of convecting branches, i.e. both partially convecting and fully convecting branches, invariant. My understanding is that the structure of convecting branches is merely determined by $M_s = M_a^*$ and γ has nothing to do with this criterion, which is different from the former two cases of varying RH_s and z_a .

d. Seasonal Cycle in the Model

In addition to convective cloud feedback, cloud albedo effect is also included into the two-layer model. Since cloud albedo effect is closely related to the solar incoming radiation at top of the atmosphere, the two-layer model should also have the seasonal variability since, for example, at polar darkness, there is no solar radiation, and the only effect of clouds is

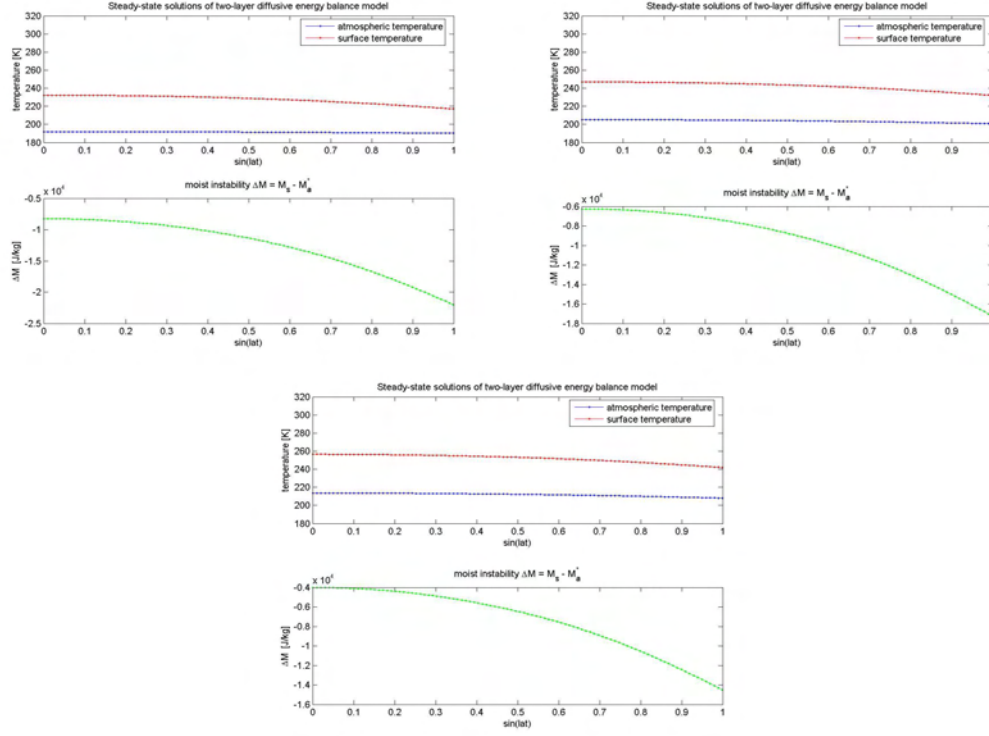


Figure 6: Steady-state surface temperature, atmospheric temperature, and moist instability measured by $\Delta M = M_s - M_a^*$ at equilibrium states $\langle 1 \rangle$ (top left), $\langle 2 \rangle$ (top right) and $\langle 3 \rangle$ (bottom) with corresponding background emissivity $\epsilon_o = 0.2, 0.6, 0.8$, respectively. Nonconvecting. Stable.

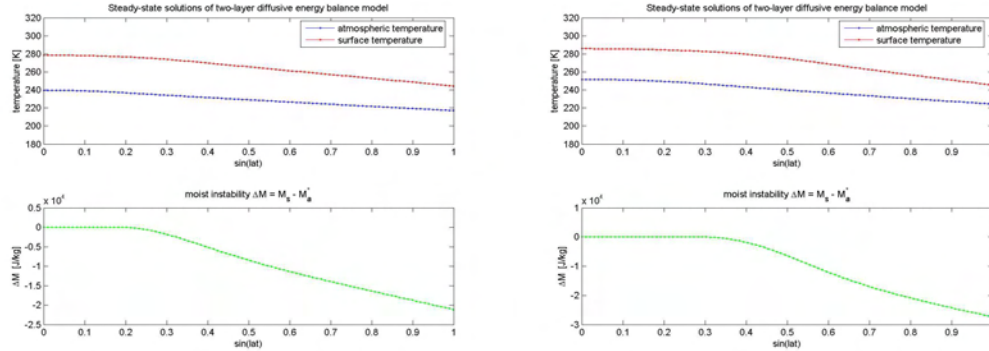


Figure 7: Steady-state surface temperature, atmospheric temperature, and moist instability measured by $\Delta M = M_s - M_a^*$ at equilibrium states $\langle 4 \rangle$ (left) and $\langle 5 \rangle$ (right) with corresponding background emissivity $\epsilon_o = 0.6, 0.46$, respectively. Partially convecting (Convection firstly takes place in tropics and extends to high latitudes, but doesn't come out at the poles yet). Unstable.

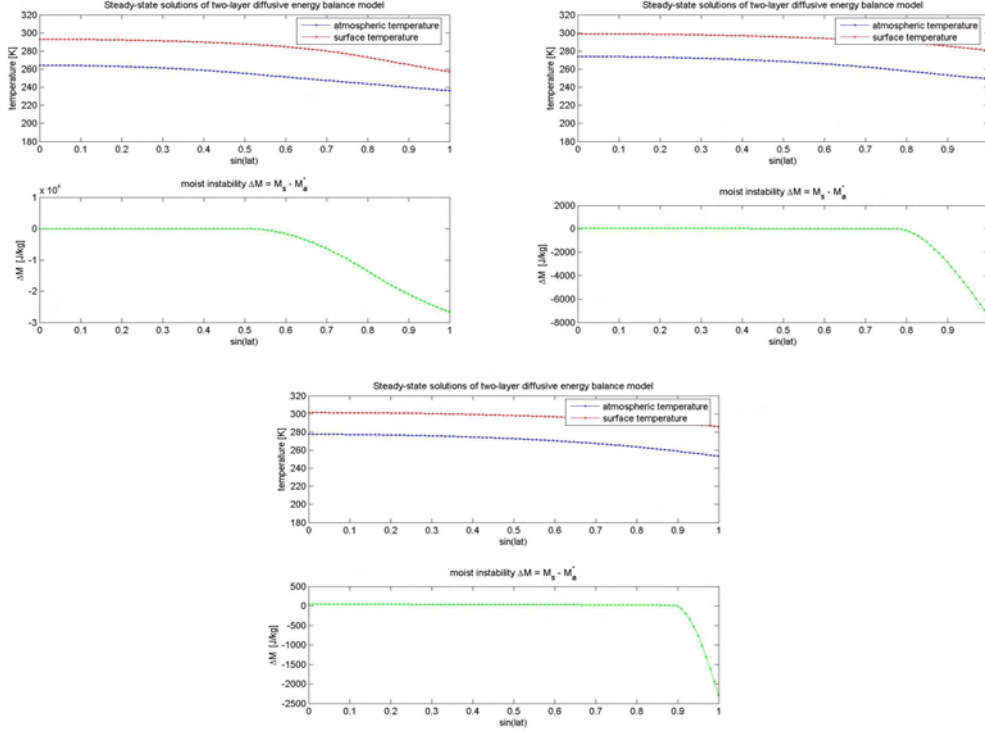


Figure 8: Steady-state surface temperature, atmospheric temperature, and moist instability measured by $\Delta M = M_s - M_a^*$ at equilibrium states < 6 > (top left), < 7 > (top right) and < 8 > (bottom) with corresponding background emissivity $\epsilon_o = 0.46, 0.6, 0.7$, respectively. Partially convecting. Stable. State < 6 > whose background emissivity $\epsilon_o = 0.46$ roughly represents the modern climate with $293K$ at the equator and $257K$ at the poles.

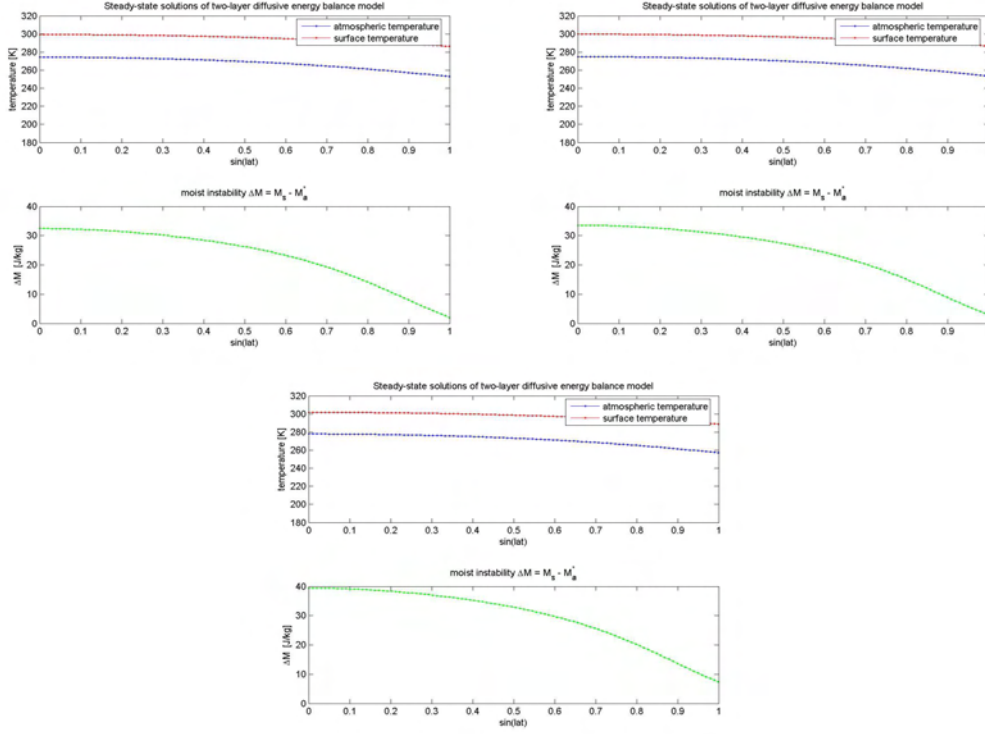


Figure 9: Steady-state surface temperature, atmospheric temperature, and moist instability measured by $\Delta M = M_s - M_a^*$ at equilibrium states $\langle 9 \rangle$ (top left), $\langle 10 \rangle$ (top right) and $\langle 11 \rangle$ (bottom) with corresponding background emissivity $\epsilon_o = 0.58, 0.6, 0.7$, respectively. Fully convecting (Convection takes place everywhere from the equator to the poles). Stable.

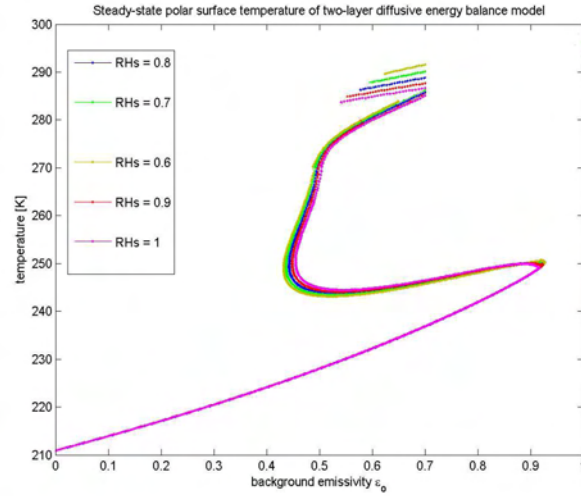


Figure 10: Steady-state polar surface temperature of two-layer diffusive energy balance model with surface relative humidity $RH_s = 0.6, 0.7, 0.8, 0.9, 1$, respectively.

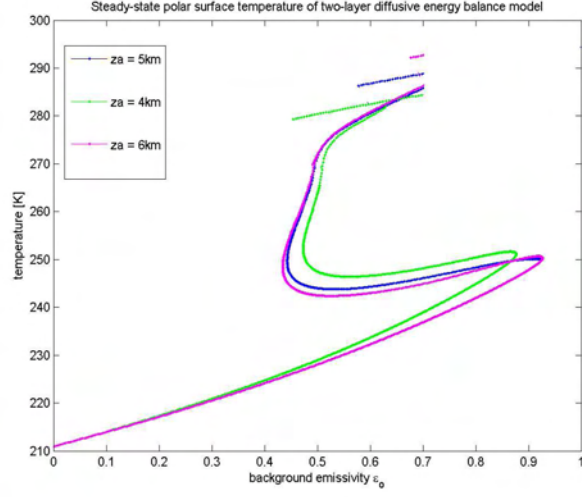


Figure 11: Steady-state polar surface temperature of two-layer diffusive energy balance model with height of the atmospheric layer $z_a = 4\text{km}$, 5km , 6km , respectively.

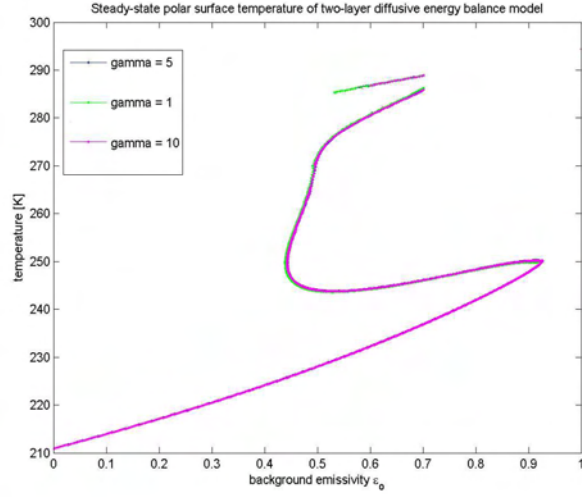


Figure 12: Steady-state polar surface temperature of two-layer diffusive energy balance model with $\gamma = 1\text{kg/m}^2/\text{s}$, $5\text{kg/m}^2/\text{s}$, $10\text{kg/m}^2/\text{s}$, respectively.

the albedo effect.

In order to figure out the average daily insolation[6] at top of the atmosphere as a function of time and latitude, some parameters are introduced beforehand. First of all, solar zenith angle θ_s is defined as the angle between the local normal to Earth's surface and the direction of incoming solar flux and determines the actual solar flux per unit surface area Q , which is given by

$$Q = L_{\odot} \left(\frac{\bar{d}}{d} \right)^2 \cos \theta_s \quad (19)$$

where \bar{d} is the mean distance from the sun where the solar constant is measured and d is the actual distance. There is less energy per unit surface area if the surface is not perpendicular to the incoming solar flux compared with that in the perpendicular case. Secondly, declination angle δ , which is defined as the latitude of the point on the surface of Earth directly under the sun at noon, varies between $+23.45^\circ$ at northern summer solstice (June 21) and -23.45° at northern winter solstice (December 21). Finally the longitude of the subsolar point relative to its position at noon is defined as the hour angle h . The relationship between the above three parameters, i.e. solar zenith angle θ_s , declination angle δ , hour angle h , and the latitude ϕ is given by

$$\cos \theta_s = \sin \phi \sin \delta + \cos \phi \cos \delta \cos h. \quad (20)$$

When it is sunrise or sunset, $\cos \theta_s = 0$, and the hour angle of sunrise and sunset is given by

$$\cos h_o = -\tan \phi \tan \delta. \quad (21)$$

Then if we plug (20) into (19) and integrate it from sunrise to sunset, the average daily insolation at top of the atmosphere can be rewritten as

$$\bar{Q}^{\text{day}} = \frac{L_{\odot}}{\pi} \left(\frac{\bar{d}}{d} \right)^2 [h_o \sin \phi \sin \delta + \cos \phi \cos \delta \sin h_o]. \quad (22)$$

However there are special conditions at the poles. When the latitude ϕ and the declination angle δ are of the same sign, which is the case in summer hemisphere, latitudes poleward of $90 - \delta$ are constantly illuminated. In addition, when the latitude ϕ and the declination angle δ are of the opposite sign, which is the case in winter hemisphere, latitudes poleward of $90 - |\delta|$ are in polar darkness. Therefore, at polar region, six months of darkness alternates with six months of sunlight. Hence the governing equations become two joint nonlinear PDEs given by

$$C_s \frac{\partial T_s}{\partial t} = \frac{\partial}{\partial y} [(1 - y^2) D_1 \frac{\partial T_s}{\partial y}] + \bar{Q}^{\text{day}} (1 - \alpha) - F_c + \epsilon \sigma T_a^4 - \sigma T_s^4 \quad (23)$$

$$C_a \frac{\partial T_a}{\partial t} = \frac{\partial}{\partial y} [(1 - y^2) D_2 \frac{\partial T_a}{\partial y}] + F_c + \epsilon \sigma (T_s^4 - 2T_a^4). \quad (24)$$

Since solar flux is not symmetric between the two hemispheres in this time-dependent case, the seasonal model and its solutions are not symmetric anymore. The condition of $T_s(y = \pm 1)$ and $T_a(y = \pm 1)$ are regular gives the boundary condition. If there is convection within some region, not only convective flux F_c and atmospheric emissivity ϵ as mentioned

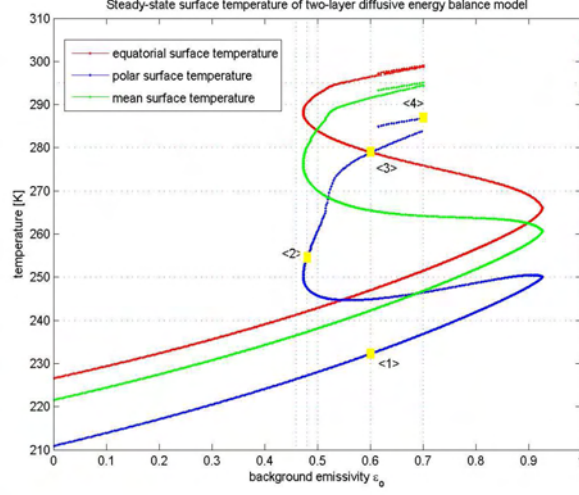


Figure 13: Steady-state surface temperature of two-layer diffusive energy balance model when cloud albedo effect is taken into account but without seasonal cycle at first. If convection takes place within some region, albedo there is assumed to increase by 0.05 in the simulation.

before, but also albedo within that region will change, for example, albedo will increase by 0.05 due to clouds' reflection of solar incoming radiation. The corresponding bifurcation diagram of steady-state solutions is represented in Figure 13 in which cloud albedo effect is taken into account but without seasonal cycle at first. It shows that due to the effect of cloud albedo, the bifurcation diagram is shifted to the right compared with Figure 5.

The above two joint nonlinear PDEs can be solved numerically and the time-dependent solutions of different initial guesses, for example, states 1 – 4 in Figure 13, are represented in Figure 14 – 17. The solutions of surface temperature at South Pole, at North Pole and at the equator are plotted every 12 months. When running the seasonal model with initial guess of state < 1 > which is a stable nonconvecting state, the time-dependent solution is wild at the beginning but eventually reaches a periodic equilibrium after about 1500 months, which is shown in Figure 14. In Figure 15, if the seasonal model starts with initial guess of state < 2 >, a stable partially convecting state, the solution quickly jumps to the snowball climate within a short time period. As shown in Figure 16, the time-dependent solution periodically oscillates around the initial steady-state equilibrium, i.e. the stable partially convecting state < 3 >. In Figure 17 which is the time-dependent solution with initial guess of state < 4 >, it is roughly periodic but with several abrupt sharp transitions. Since these abrupt transitions are transient, the intermediate states are unstable. It is interesting to notice that state < 4 > is a stable fully convecting state and is very close to the unstable missing branch that we mentioned before. It is reasonable to explain the existence of these abrupt transient transitions in this way: oscillations around the initial steady-state equilibrium < 4 > allow the possibility of transition to lower unstable branch, however, the time-dependent solution has to find other stable states such as the stable partially convecting states and the stable fully convecting states due to the instability of

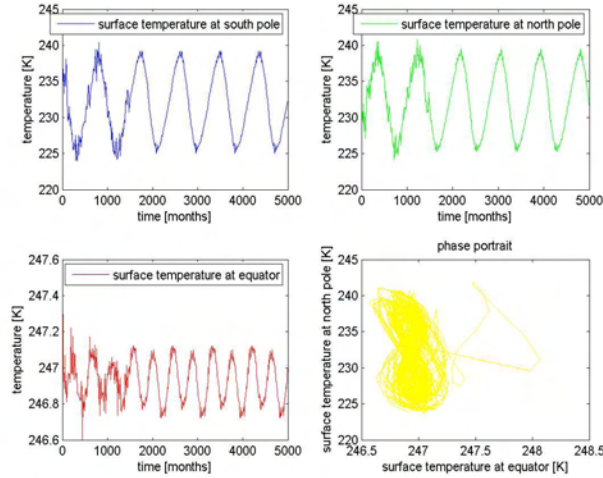


Figure 14: Time-dependent solution of the seasonal model with an initial guess of state $< 1 >$ in Figure 13, i.e. $233K$ at the poles and $247K$ at the equator and with background emissivity $\epsilon_o = 0.6$. Time is in the unit of month. The top left panel is the surface temperature at south pole as a function of time, the top right one is the surface temperature at north pole as a function of time, the bottom left one is the surface temperature at the equator as a function of time, and the bottom right one is the phase portrait with x-axis the surface temperature at the equator and y-axis the surface temperature at north pole.

the lower branch. It is another indirect indicator of the existence of the unstable branch that is related to the transition of polar convection.

Conclusions

We start the problem trying to figure out the effect of water vapor in the climate system in simplifies model simulations. A one-layer diffusive moist static energy balance model is applied at first. From the model output, three possible steady climate states, i.e. a stable snowball climate, an unstable climate in which surface temperature is about $20K$ lower than the modern value everywhere, and a stable ice-free climate, have been found, moreover, the effect of water vapor is represented mainly by the meridional moisture transport to the poles through the comparison with dry model. However, even if all the model parameters are based on modern climate observations, the one-layer model is unable to produce the modern climate. In order to better simulate the real climate, a two-layer diffusive energy balance model has been developed in which an important effect of water vapor, i.e. convective cloud feedback, is parameterized through the impact of convective flux and atmospheric emissivity. The solutions of this model are very interesting. First of all, it is able to roughly represent the modern climate which in turn indicates the significance of clouds effect in real climate. Furthermore, two more steady equilibrium states, besides the three equilibria turned out in one-layer model, have been found due to the implement of convective cloud feedback. Significantly, each equilibrium has its clear physical meaning. For example, when

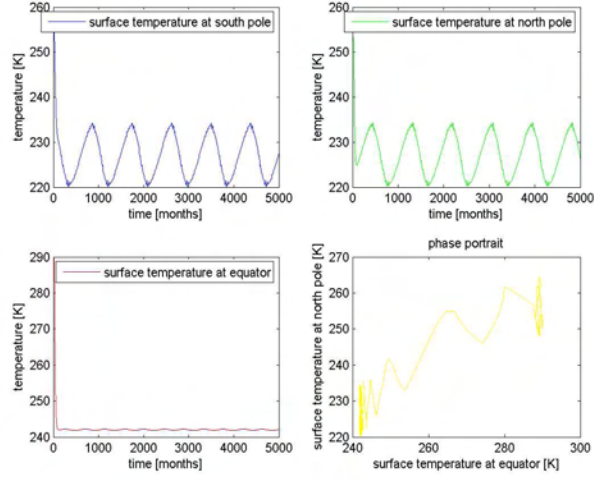


Figure 15: Time-dependent solution of the seasonal model with an initial guess of state $\langle 2 \rangle$ in Figure 13, i.e. $254K$ at the poles and $290K$ at the equator and with background emissivity $\epsilon_o = 0.48$. Time is in the unit of month. Others are the same as Figure 14.

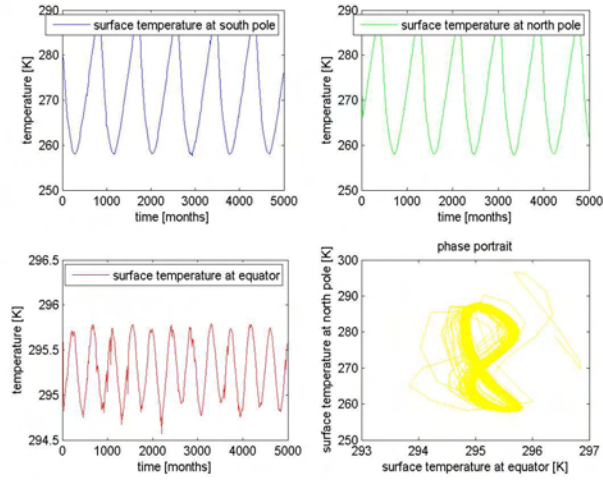


Figure 16: Time-dependent solution of the seasonal model with an initial start of state $\langle 3 \rangle$ in Figure 13, i.e. $279K$ at the poles and $296K$ at the equator and with background emissivity $\epsilon_o = 0.6$. Time is in the unit of month. Others are the same as Figure 14.

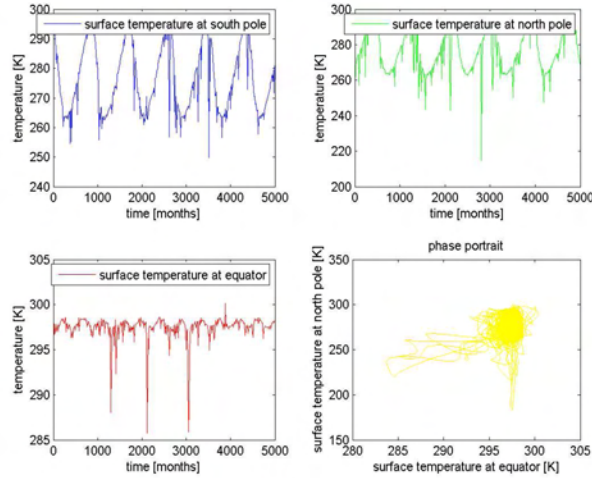


Figure 17: Time-dependent solution of the seasonal model with an initial start of state $< 4 >$ in Figure 13, i.e. $287K$ at the poles and $299K$ at the equator and background emissivity $\epsilon_o = 0.7$. Time is in the unit of month. Others are the same as Figure 14.

surface temperature is very low everywhere, there is no possibility of convection because it is always moist stable, and it is the same equilibrium state that is noticed in one-layer model. As the equatorial surface temperature reaches some critical value, convection takes place at the equator firstly. As the equatorial surface temperature gets higher, convection takes place at high latitudes. Eventually there is convection at the poles when temperature is relatively high. One of the two more equilibria in two-layer model represents a stable climate state with convection taking place everywhere from the equator to the poles while the other one is an unstable branch that is related to the transition of polar convection. The last section of the project is the seasonal model, in which both the seasonal cycle of solar insolation and cloud albedo effect are included. The time-dependent solutions are basically oscillations around steady equilibria of constant solar insolation. Depending on how large the variability of climate is at each state, it is possible to switch from one equilibrium to another, for example, the transition to snowball climate from a warmer climate state.

Although this study is basically based on a simple two-layer energy balance model, it is able to capture the basic but essential physics of convective cloud feedback and produce multiple steady equilibria in addition to temperature distribution of modern climate. Since the model is a simplified version, the physics can be understood completely and clearly. Furthermore, it indicates the importance of convective cloud feedback of real climate. However, since the model is simply parameterized and only includes few effects and feedbacks in real climate, the results from the model cannot be simply compared with real climate situation. For example, albedo is a complicated variable which not only depends sensitively on the surface and clouds situation but also varies with time and has feedback to the climate, and the real effect of clouds is much more complicated than what is parameterized in the model. More numerical computation and analysis will be done on the seasonal model and its time-dependent solutions in the future.

Acknowledgements

Great thanks to Neil Balmforth, who helped me throughout the whole summer about the project, Eli Tziperman, who gave me helpful suggestions on the improvement of models and physical interpretation, and Raymond Pierrehumbert, who brought up this interesting project. Also many thanks to Joseph Keller and Dorian Abbot for helpful discussions. Final thanks to the Geophysical Fluid Dynamics summer program and all the faculty and fellows.

References

- [1] D. S. ABBOT AND E. TZIPERMAN, *Controls on the activation and strength of a high latitude convective cloud feedback*, J. Atmos. Sci., (2008).
- [2] ———, *A high latitude convective cloud feedback and equable climates*, Quart. J. Roy. Meteor. Soc., 134 (2008), pp. 165–185.
- [3] ———, *Sea ice, high latitude convection, and equable climates*, Geophys. Res. Lett., 35 (2008), pp. L03702, doi:10.1029/2007GL032286.
- [4] M. I. BUDYKO, *The effect of solar radiation variations on the climate of the earth*, Tellus, 21 (1969), pp. 611–619.
- [5] P. Z.-G. DARGAN M. W. FRIERSON, ISAAC M. HELD, *A gray-radiation aquaplanet moist gcm. part 2: Energy transports in altered climates*, J. Atmos. Sci., 64 (2006), pp. 1680–1693.
- [6] D. L. HARTMANN, *Global Physical Climatology*, Academic Press, 1994.
- [7] IPCC, *Climate Change 2007, Synthesis Report to the fourth assessment of the Intergovernmental Panel on Climate Change*, Intergovernmental Panel on Climate Change, 2007.
- [8] G. R. NORTH, *Theory of energy-balance climate models*, J. Atmos. Sci., 32 (1975), pp. 2033–3043.
- [9] G. R. NORTH, R. F. CAHALAN, AND J. JAMES A. COAKLEY, *Energy balance climate models*, Reviews of Geophysics and Space Physics, 19 (1981), pp. 91–121.
- [10] T. SASAMORI, *The radiative cooling calculation for application to general circulation experiments*, J. Appl. Meteor., 7 (1968), pp. 721–729.
- [11] W. D. SELLERS, *A climate model based on the energy balance of the earth-atmosphere system*, J. Appl. Meteor., 8 (1969), pp. 392–400.

